Antitrust Chronicle

Data, Data Everywhere...

TABLE OF CONTENTS

LETTER FROM THE EDITOR

Dear Readers,

In this Chronicle we turn to the issue of the interaction between big data and competition. Is big data a game changer?

In recent years, big data has variously been discussed as the "new oil," a barrier to entry to digital markets (and even brick and mortar sectors), or alternatively as a mere asset (albeit a new one), like any other.

Yet no clear consensus has emerged. Do large datasets justify modulation of existing antitrust doctrines? Are they better regulated through bespoke regulation? Or are the existing rules sufficient? How would any of the mooted changes interact with the burgeoning growth of personal data privacy regulation?

These questions, among others, have been raised in academic discussion, discourse among practitioners, and in various antitrust decisions in the past years. Yet there are, to date, no clear answers.

Whatever they may be, these answers will be key to the enforcement of competition rules for years to come. We hope this Chronicle provides a window into this fascinating debate.

Lastly, please take the opportunity to visit the CPI website and listen to our selection of Chronicle articles in audio form from such esteemed authors as Maureen Ohlhausen, Herbert Hovenkamp, Richard Gilbert, Nicholas Banasevic, Randal Picker, Giorgio Monti, Alison Jones, and William Kovacic among others. This is a convenient way for our readers to keep up with our recent and past articles on the go, in the gym, or at the beach.

As always, thank you to our great panel of authors.

Sincerely,

CPI Team

Scan to Stay Connected!
Scan or click here to sign up for
CPI's **FREE** daily newsletter.

SUMMARIES

07

CPI Talks…

…with Paul Gilbert & Maurits Dolmans

In this month's edition of CPI Talks we have the pleasure of speaking with Paul Gilbert and Maurits Dolmans, of the London office of Cleary, Gottlieb, Steen and Hamilton LLP. Maurits recently wrote articles entitled "Should We Disrupt Antitrust Law?" and "Pandora's Box of Online Ills." Paul has given presentations on "Competition Law and Big Data."
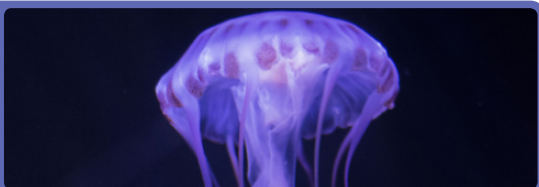
11

Digital Data as an Essential Facility: Control

By Catherine Tucker

This paper examines the question of data as an essential facility. It argues that rather than size being the biggest driver of data meeting the definition of being an essential facility, it is instead whether the data are protected by copyright or privacy regulation.

16

The Impulse to Condemn the Strange: Assessing Big Data in Antitrust

By Alexander Krzepicki, Joshua D. Wright & John M. Yun

In current antitrust policy debates, it is almost a foregone conclusion that digital platforms' collection and use of "big data" is a barrier to entry. In this article, we argue that big data should properly be considered a two-stage process. The reason why this classification matters is because it allows us to link big data to concepts that antitrust is already familiar with: economies of scale, learning by doing, and research & development. By linking big data with the familiar, we hope to avoid a common tendency in antitrust to condemn the strange.

21

Essential Facilities Fallacy: Big Tech, Winner-Take-All Markets, and Anticompetitive Effects

By John Pecman, Paul A. Johnson & Justine Reisler

There have been calls to treat data as an essential facility to reduce barriers to entry in the digital economy where a "winner-take-all" dynamic is often present. In such markets, traditional enforcement may fail to enjoin conduct, or an acquisition, that lowers welfare in expectation because anticompetitive effects, regardless of their magnitude, are given no weight if deemed unlikely to occur. The authors consider several proposals to fine-tune competition law and argue that calls to regulate Big Tech companies like public utilities should be viewed with skepticism because they ignore dynamic competition, the complexity of micromanaging a rapidly evolving sector of the economy, and the very purpose of competition law – to protect competition, not competitors.

SUMMARIES

31

Big Data, Big Target for EU Antitrust Enforcement?

By Jay Modrall

In her new five-year mandate as Competition Commissioner, Commission Executive Vice-President Vestager will be considering changes to EU antitrust enforcement in relation to big data in several respects: the role of data in defining markets and assessing market power; imposing data-access remedies in abuse-of-dominance cases; and clarifying rules on sharing and pooling data. In merger review, pressure for reform is coming from many directions, but (perhaps surprisingly) not particularly in relation to the assessment of big data combinations. Based on the May 2019 special advisor report, the most significant legal changes may come in the form of a new and less stringent approach to data access remedies, but such cases are few and far between and often take years to resolve. In parallel, Vestager will be working on new regulatory frameworks for artificial intelligence and digital services. Vestager's views on the potential for big data to confer market power and the pros and cons of mandating access to data may ultimately have a greater role in shaping EU regulation than EU antitrust enforcement.

37

Can Digital Data be Replaced? Data Substitutability is the Key

By Mariateresa Maggiolino & Giulia Ferrari

The question of whether digital data constitute a barrier to entry or an essential facility does not admit a "one-size-fits-all" answer: only a case-by-case assessment of whether digital data have viable substitutes will allow the antitrust law interpreter understanding whether digital data fall within one of the categories. Therefore, the real question becomes "do digital data have viable substitutes?" In order to answer that question, a thorough assessment of the needs and uses that the undertaking seeking those digital data wants to accomplish through them is required. Establishing the kind of information that the undertaking wants to infer from the digital data, indeed, will be crucial in order to understand whether such "demand" can be met by recurring to potential alternative datasets. However, the Fourth Industrial Revolution has also revealed competitive risks that the categories of traditional antitrust analysis do not seem to tackle effectively.

WHAT'S NEXT?

For March 2020, we will feature Chronicles focused on issues related to (1) **LeadershIP EU**; and (2) **China Edition**.

ANNOUNCEMENTS

CPI wants to hear from our subscribers. In 2020, we will be reaching out to members of our community for your feedback and ideas. Let us know what you want (or don't want) to see, at: antitrustchronicle@competitionpolicyinternational.com.

CPI ANTITRUST CHRONICLES APRIL 2020

For April 2020, we will feature Chronicles focused on issues related to (1) **Sports**; and (2) **Remedies**.

Contributions to the Antitrust Chronicle are about 2,500 – 4,000 words long. They should be lightly cited and not be written as long law-review articles with many in-depth footnotes. As with all CPI publications, articles for the CPI Antitrust Chronicle should be written clearly and with the reader always in mind.

Interested authors should send their contributions to Sam Sadden (ssadden@competitionpolicyinternational.com) with the subject line "Antitrust Chronicle," a short bio and picture(s) of the author(s).

The CPI Editorial Team will evaluate all submissions and will publish the best papers. Authors can submit papers on any topic related to competition and regulation, however, priority will be given to articles addressing the abovementioned topics. Co-authors are always welcome.

CPI TALKS…

…with Paul Gilbert & Maurits Dolmans

In this month's edition of CPI Talks we have the pleasure of speaking with Paul Gilbert and Maurits Dolmans, of the London office of Cleary Gottlieb, Steen and Hamilton LLP. Maurits recently wrote articles entitled "Should We Disrupt Antitrust Law?" and "Pandora's Box of Online Ills." Paul has given presentations on "Competition Law and Big Data." Both have extensive experience in the IT sector. They have advised various clients in this area, but speak here on their own behalf.

Thank you, Paul and Maurits, for sharing your time for this interview with CPI.

1. **Recent months and years have seen intense discussion on "big data" and its role in competition in the digital economy. Much of this discussion relates to the question of whether, in the language of the competition rules, "big data" are a mere asset, a "barrier to entry," an "essential facility," or all of the above. Is the classic terminology misleading when applied to this new phenomenon? Rather, does this debate merely underline the need for a careful case-by-case approach to competition issues involving complex technological questions?**

We don't think the terms need changing. They are part of a framework that is used to differentiate those situations where competition intervention is needed from those where it is not. Intervening in the wrong cases will only undermine incentives to invest in collecting and analyzing data.

In most cases, data are an input for a product or service. It is not the service itself. So, is it right to ask whether data are an essential facility? Put another way, do competitors need access to that particular set of data for competition to exist in a downstream market? If not, intervention is unwarranted, and could do more harm than good. There may, in theory, be cases where data are an essential facility, because the data are indispensable and the source is no longer available. But they will be the exception rather than the rule.

Holding data is not like owning a harbor or controlling the electricity network. In those cases, anyone trying to operate a shipping service or supply electricity may have to have access to the infrastructure. Data are not "owned" or controlled in that way. I could record the color preference of every car buyer if I thought the information was going to be useful, and I may be only person to "own" that data. But that doesn't make that database an essential facility, and it doesn't stop anyone else doing exactly the same thing. Many types of data are "non-rivalrous goods," in the sense that they can be duplicated and recreated without limit.

It is also tempting to treat all types of data as the same, when they are not. Personal data are different from observed data, for instance. Observed data often has a short lifespan. For my car-color database to remain useful, I would probably have to repeat the exercise regularly. Otherwise the data will become stale and worthless. And I may not collect it at all if I had to give it away to all my competitors.

Personal data are different. If a consumer uploads their photographs and contacts to a social media platform, they could become locked in. If many users are locked in, that may create a barrier to entry for new social media platforms – in the sense that gathering an equivalent set of data is going to be more difficult or more expensive for the new entrant than it was for the first mover. In such a situation, ensuring that consumers can transfer their information freely and easily between different platforms can be pro-competitive. This, in turn, may require standard setting to ensure that data can be easily ported to new service providers.

So, the terminology is right. It just needs to be applied carefully, and seen in context.

2. Information scientists famously refer to the so-called "DIKW pyramid." In short, mere raw "data" provide "information" that experts can use to produce "knowledge" and (hopefully) "wisdom" that can yield valuable results. This leads some commentators to suggest that, at least in certain circumstances, data should not be considered to be a true "barrier to entry": the real barrier would be the technology and expertise needed to extract "wisdom." To what extent should competition practitioners, enforcers, and courts rely on expertise from the sciences to assess these and other questions in individual cases?

We agree. In this information pyramid, data are merely raw measurements, whereas information is an understanding of the relations between the data, knowledge is an understanding of the patterns that emerge, and intelligence or wisdom is an understanding of the principles that allow judgment, better decision making, and prediction.

Data are useless without the ability to analyze them intelligently and creatively. Take away the data from any online service provider (but leave the engineers) and they will quickly be back in business. Take away the engineers (but leave the data) and they will soon be in trouble.

There are lots of examples. Start-up companies such as WhatsApp, Hailo, BlaBlaCar, Snapchat, Instagram and Pinterest have all been able to build successful and innovative products with access to little or no data at the start. If data are a barrier, it can certainly be overcome.

Studies have shown that data quickly loses its incremental value after a certain point. According to the law of diminishing returns, once a critical volume is reached, the marginal value of additional data for statistical analysis is limited. There is no value in simply collecting more and more data.

Competition authorities and courts should take this into account. They need to ask what data is really needed, how much is needed, and how easily it can be collected. Data is everywhere, it is cheap to collect, can be bought and sold, and is non-rivalrous – it is not used up like oil.

So, it is right to think of intelligence and creativity as the real drivers of competition. As to whether they could become barriers to entry, it seems unlikely. There seems to be no shortage of either.

3. Recent reports (including the UK Furman Report, the EU Commission's Crémer Report, and the ACCC's Digital Platforms Report) discuss the treatment of data under the competition rules in some detail. Specifically, the Crémer Report underlines the need to adopt a careful approach to the "essential facilities" concept if applied to data. In your view, are there grounds for a modified approach to the "indispensability" and "new product" criteria under the classic essential facilities rules should they be applied to data?

Many competition authorities and regulators are grappling with the question whether we should be quicker to intervene in digital, data-driven markets. Should companies that collect data be required to give access to that data as a way of helping their competitors, even if they have not done anything that would otherwise be considered anticompetitive?

This is market engineering rather than antitrust, and is a risky path to follow. It means forcing Firm A to support Firm B even if Firm B does not need that support to compete. Forcing data-sharing in this way would pervert the normal competitive incentives of both. Firm A will be less inclined to invest in collecting data and designing innovative ways of doing so if it has to share the results with its competitors. Firm B will not have incentives to invest either, because it can free-ride on the efforts of Firm A.

The Crémer Report recognizes precisely this concern. It recognizes the risks that a lower threshold for intervention would have on commercial incentives. It says that "*a thorough analysis will be required*" to determine whether access to data is "*truly indispensable*" in order to compete in neighboring markets. It also recommends that competition law principles should be central to any sector-specific regulation that is designed to mandate data access.

We agree. It is easy to lose sight of the benefits that come from companies investing in collecting and analyzing data in ever more inventive ways. Undermining these incentives could be hugely damaging. As we said before, with respect to personal data, the better option is to give individuals the right to port their data between different services providers, so as to encourage the latter to keep improving their offering to persuade the users to stay with them.

4. One of the most controversial forms of "big data" relates to personal information used to provide services such as online social media, search, and advertising. Individuals' rights concerning such data are regulated separately under rules such as the EU GDPR and the California Consumer Privacy Act. Certain competition remedies (notably mandated access) would raise obvious potential conflicts with these specific regulations. How should this circle be squared? Is this an argument for non-intervention by competition enforcers in cases of such potential conflict, or at least careful remedy design (e.g. anonymization of user data if disclosure is mandated)? Is enough being done by data protection regulators and competition enforcers to coordinate in this regard?

It would be perverse for competition law remedies to override data protection. Central to GDPR and other data protection rules is the principle that individuals should be able to control what happens to their personal data – information about them.

I may be happy to share my personal information, health records and banks details with one company that I trust, but not with others. That is undermined if the company is then forced to share my details with someone else that I don't trust, or have perhaps never heard of. When a consumer shares their personal data, they have to be confident that it will be protected and won't be shared without their consent.

A good example of how the system can work is in Open Banking. If I want to use a new service – perhaps a FinTech start-up company – that company will be able to access my bank account details securely, but only if I give my consent.

Many of the largest technology companies (including Apple, Facebook, Google, Microsoft and Twitter) are currently working on a similar initiative, called the Data Transfer Project. The purpose is to allow consumers to transfer their information freely between providers using open source coding. It will mean that users do not have to download their data from one service before uploading to another. This is procompetitive. It helps users to switch even more quickly between providers and try out new services, as well as backing up their data on multiple services. And the user always remains in control. This approach to "portable data" creates new opportunities for intermediary services or portals that help users transfer data, or manage multiple platform settings.

This type of initiative facilitates competition and new entry, while still respecting consumers' rights and data protection rules.

5. A common complaint levied against companies possessing large datasets is that their existing "scale" renders it impossible for new entrants to compete. Are there alternative means that regulators, policymakers, and industry could use to make large-scale data available to new entrants (e.g. databases of photo, video, or anonymized personal data that could be used to train machine learning algorithms)? Would such an approach be feasible as a remedy in a hypothetical competition case? Or is such an approach best left to bespoke industry or government-mandated initiatives?

It is wrong to assume that scale is necessarily critical or a barrier to entry. This idea ignores the law of diminishing returns. Experience in many areas where data are used as an input, like machine translation, image processing, speech recognition, or teaching artificial intelligence systems, the first million data are much more important than the second and so forth. Huge amounts of open source datasets are increasingly available for AI training (from sources as various as kaggle.com, FiveThirtyEight, Google Public Data Explorer, or from EU Open Data Portal, Data.gov. WHO Open Data repository, and World Bank Open Data, etc.). Open source software and cloud computing makes it easier for start-ups than ever before to run complex computations on these open data.

We have already mentioned several examples of companies that have been able to grow quickly without having access to large datasets when they started. In each of these cases, "scale" was not a barrier. If anything, they show that innovation is critical and that data are easy to gather.

The European Commission has also considered this question many times in merger cases. Whenever it has looked at markets in detail, it has found that rivals are able collect the data they need. In *Facebook/WhatsApp*, for example, it found that "*there will continue to be a large amount of Internet user data that are valuable for advertising purposes.*"

In *Microsoft/Skype*, it found that barriers to entry were low, pointing to the "*immediate success*" of new entrants Viber, Fring, and Tango. Viber was downloaded more than a million times within three days of its launch, 10 million times within two months, and 15 million times within six months.

There may be cases where access to an existing database is critical for competition to exist. It will depend on the type of data and whether the database is replicable. In the UK, for example, the National Health Service provides access to huge amount of anonymized healthcare data for analysis. That data wouldn't be available otherwise and cannot be replicated. But this is an exception. Online user data are very unlikely to fall into this category.

In the EU, we have a tendency to blame U.S.-based online firms (ignoring China for the time being) for the lack of EU online powerhouses, and for the difficulties of digital disruption. They are conveniently big targets. But a focus on data sharing is not the solution. The societal problems resulting from digital disruption should be solved by targeted and proportional regulation, rather than forcing firms to share data across the board. If we want to grow EU-based innovative powerhouses, the answer is a policy of creating an environment where innovation can flourish: innovation hubs combining academic centers of excellence and existing technology businesses with carefully targeted Government-sponsored open projects, which together can fertilize start-up businesses by providing talent and risk capital, a culture of entrepreneurship, critical thinking, and creativity, nurtured in an environment of fair and balanced tax laws, IPRs strong enough to encourage innovation but not so strict as to block new entry, and competition rules that preserve opportunities for new entrants without punishing them once they become successful. There is plenty of big data; what is scarcer is technical talent and business creativity, and wise regulators.

DIGITAL DATA AS AN ESSENTIAL FACILITY: CONTROL

BY CATHERINE TUCKER[1]

1 Catherine Tucker is the Sloan Distinguished Professor of Management Science and Research Associate at the NBER. Please see my disclosure statement at https://mitmgmt-faculty.mit.edu/cetucker/disclosure/.

I. ARE DATA AN ESSENTIAL FACILITY?

The notion of an "Essential facility" is key to antitrust law, because denial of access to key facilities can mean that a potential monopolist will be immune, at least for some time, to most forms of competition. As an economist, I observe that legal scholars debate whether anything can really be an essential facility, given that one can always raise exceptions or caveats to any even potentially watertight case.

Given this scholarly debate, it is useful to introduce terminology used for a parallel notion in strategic management. Here the focus is on a "resource" that gives "sustainable competitive advantage," or that can act as a barrier to entry. The idea of such a "resource" was first articulated by my colleague Birger Wirnerfelt at MIT,[2] and then further refined by Barney (1991).[3] To be a "resource," the asset in question needs to be valuable, non-imitable (or at least difficult to substitute), and rare. More recent theory has introduced the additional nuance that the firm must be able to "control" this rare, valuable, imperfectly imitable and non-substitutable resource.[4]

When we teach this framework, our aim is to give managers a means to consider what resources a firm has, which may allow them to build sustainable competitive advantage in such a way that would allow them to profit. However, a parallel use of this framework can also be to allow an examination of potential barriers to entry, and, at the extreme, whether or not a firm's "resource" or "core competency" has morphed to the extent that it could be considered to be an essential facility.

In earlier work I went through the first four pillars of this strategy framework (i.e. the conditions relating to a given facility of asset being rare, valuable, non-imitable and non-substitutable) in considering whether data could confer such an enduring competitive advantage.[5] In this essay, I will consider the more recent additional pillar that has been added to this strategic framework – that is, the question of whether firms really control data in a way that makes it a source of competitive advantage.

II. INTERNAL MEANS OF CONTROL OVER DATA

The type of data that firms can directly control access to differs according to its purposes and between firms. Typically, the key distinction is whether the presentation of data is internal or external. For example, YouTube cannot easily prevent access to data about who has liked a video on YouTube, or the content of comments. However, Uber can choose not to make public data on the geographic granularity of rider pickups. Of course, even in that case it is somewhat possible to recreate the data, and indeed a French PhD student has managed to access partial data through the Uber API for her work on gender discrimination in ride pickups.[6] In general, though data is non-rivalrous, it is possible to exclude access to particular data if the data are not public. Sometimes, the legal treatment of data has focused on the idea of non-rivalry – which is indeed a key component of the definition of a public good – without also acknowledging that much of the time the same digital tools that allow the collection of vast datasets also permit control over who accesses it.

The key point is that firms are often able to control access to a particular dataset. What they are far less able to do is to control the ability of rival firms to create a similar dataset. In particular, they are unable to control the ability of a rival firm to create a dataset which offers similar insights. This latter point is important because ultimately the value of data is not the raw manifestation of the data itself, but the ability of a firm to use this data as an input to insight.

In general, I argue that the ability of a firm to control whether a rival also creates a similar dataset very much depends on the extensiveness of the digital footprint that a consumer has when generating data that gives that particular insight.

Let us contrast a few scenarios which are all taken from the context of digital advertising. I recognize that the power of digital data may well be greatest outside the world of digital advertising (for example in optimizing logistics). I focus on digital advertising simply because that is the focus of much of the current antitrust debate involving data in this context:

---

2 Wernerfelt, Birger, "A resource-based view of the firm," Strategic Management Journal 5.2 (1984): 171- 180. The original article has been cited over 33,000 times on Google scholar. It was also popularized by Prahalad, C. K. & G. Hamel, "The core competence of the corporation," Harvard Business Review 68.3 (1990): 79-91.

3 Barney, Jay B., "Firm Resources and Sustained Competitive Advantage," Journal of Management 17.1 (1991): 99-120. This paper has over 70,000 Google Scholar citations.

4 Barney, Jay, Mike Wright & David J. Ketchen Jr., "The resource-based view of the firm: Ten years after 1991," Journal of Management 27.6 (2001): 625-641.

5 Lambrecht, Anja & Catherine E. Tucker. "Can Big Data protect a firm from competition?," Competition Policy International (2015).

6 "Algorithm Bias on Uber" – Clara Jean, Université Paris Sud, Working Paper (2020).

1. A firm owns data allowing them to identify both the income and zip code of a particular set of "eyeballs" arriving at a website, and thereby to know what hotel ad to show to that set of eyeballs.[7]

2. .A firm owns data that allows them to predict what smartphone a set of "eyeballs" is most likely to buy, and to feature this smartphone in personalized recommendations.

3. A firm owns data that identifies that a set of "eyeballs" is shopping for last-minute flowers to be delivered on Mother's Day.

As you read this essay, it might be useful for you to rank these examples in terms of likelihood of the firm being able to "control" this data as a source of sustained competitive advantage. I would argue that in each instance, there is a feature of the data which limits control.

A. When Data Reproducibility Limits Control

In the Stigler report, a motivating example for why data might have increasing returns is the idea that one potential source of data about hotel bookings is locational data.[8] The second useful source of information might be presumed income of the individual. The Stigler report argues that the combination of data about zip code and income may be more powerful than the information about each data point individually. However, what is striking about this example is that the two sources of data (zip code and income) are publicly available, and are therefore not hard for any firm to gain access to.

The fact that data is non-rival and virtually costless to produce has led to a large industry of data brokers such as Acxiom and Experian, who collect and parse data about people's internet activity. This is usually the focus of discussion about data aggregation and reuse. However, this example of location and income information is a useful example of the types of data which are freely available through public datasets. It is reasonably easy to geo-locate any set of eyeballs using a digital device.[9] It is also reasonably easy to use census data to obtain the average income of a zip code.[10]

Given how relatively easy it is to reproduce zip code and income information from public sources and then link them to the geographical location of a user, then this will inherently limit a firm's ability to control such data as a source of sustained competitive advantage.

B. When the Ability to Recreate the Insights of the Data Limits Control

In the second example, it is notable that the potential source of competitive advantage is not data *per se*. Instead, it is the prediction based on data about what smartphone someone is likely to purchase. Indeed, no company knows for certain what smartphone someone will buy.[11] There are many potential ways to predict smartphone purchase likelihood.

One potential course is to use paid search data on the type of phones someone has searched for. Another way might be to use browsing data on a review site to see what kind of reviews the user is browsing. Another way may be to use "look-a-like" data to study what smartphones other similar people have bought in the past.

---

7 This is a scenario discussed in the Stigler Committee on Digital Platforms Final Report (2019), https://www.publicknowledge.org/wp-content/uploads/2019/09/Stigler-Committee-on-Digital-Platforms-Final-Report.pdf.

8 In the Stigler report it states that the "home zip code" of the user is useful as there may be some "limited ability to predict interest in a hotel based on this zip code being sufficiently far from the hotel." The report formally models the likelihood of a user booking a hotel as a function of their distance from the zip code. I assume this means this implies that I am most likely by their model to book a hotel in Augusta (Western Australia), which is the city I believe furthest from Boston. Interestingly, before reading the Stigler report prediction that I would book a hotel in Augusta, I had not contemplated going to Augusta. I agree it does look like a nice destination, though given that Augusta only has two hotels, I suspect that the prediction task is not that hard.

9 For example, one can access this information for browsers from https://www.w3schools.com/ html/html5_geolocation.asp or via other platforms such as https://developers.google.com/web/ fundamentals/native-hardware/user-location. There are of course issues with accuracy if people take steps to deliberately mask their location, but these are likely to be shared by all firms.

10 Go to https://catalog.data.gov/dataset/zip-code-data.

11 I am not sure what smartphone I will buy next.

One issue is that this kind of prediction task is not easy. I have studied this in my own research – (Neumann et al., 2019) – where we show that often predictions of people's age and gender based on browsing data are not much better than random chance.[12] However, that research seems to suggest that it is the skill of the humans in charge of the predictions that seems to matter, in that prediction improves when a human guides the choice of the algorithms.

Therefore, the key questions are, whether there is only one set of data that is appropriate for a prediction task; and also to give appropriate credit for the difficulty of creating insights from the data, rather than ascribing the rarity of the insight to the data itself.

C. When the Short-lived Nature of the Data Limits Control

I chose the example of the Mother's Day flowers, because this is an instance where the consumer is likely to have a small digital footprint. Someone running late for Mother's Day is likely to perhaps only engage with one search engine or one website when searching for flowers. As a result, it is possible that only one firm has access to the insight that someone is wishing to buy flowers for Mother's Day. This enhances a firm's potential control of the data.

However, there is a caveat, which is that in this particular example the value of the data is very short-lived. Once Mother's Day is over, the data loses its value, rendering it not a good source of sustained competitive advantage. Next Mother's Day, the firm will have to compete again to be the digital platform or website which gains the insight. The only chance of the data retaining its value is that the same person may be tardy at buying flowers on Mother's Day the next year, such that the data become useful as an annual prompt.

This example shows that it is hard to generalize about how easy it is to `control' the value of any one bit of data. However, it also shows that it is important when thinking about control of data, to also think about how long-lived any insight from the data might be.

III. EXTERNAL SOURCES OF CONTROL: PRIVACY REGULATION AND COPYRIGHT AS SOURCE OF CONTROL OVER DATA

The previous section discussed limitations in terms of imposing internal control over data. However, I would argue that there are still concerns regarding external sources of control. I would argue that this is most likely to be an issue when the data in question is also potentially governed by privacy or copyright protections.

A useful example where data segmentation services are treated as though it was an essential facility is that of IMS Health and NSC Health corporations. These are two competitors in pharmaceutical data services in Germany which provided sales reports from individual pharmacies.

Due to German privacy laws, data has to be appropriately anonymized in a way which is privacy-compliant. The most practical way of doing this (without losing all marketing insight from the data) is to aggregate it to the postcode level. This was done by something referred to as the "brick structure," which grouped together pharmacies into commercially useful geographical clusters which would not permit the identification of any one pharmacy and complied with the German privacy rule that at least 3 pharmacies had to be aggregated. However, IMS asserted copyright over the brick structure. Any other means of aggregation were potentially not privacy-compliant, because differential aggregation could in theory then be used to identify an individual pharmacy if a firm cross-referenced the two datasets. At this point, the European Commission ordered IMS to grant access to the brick structure on commercially reasonable terms.

It is notable though that in this case it is not data *per se* that is the essential facility but instead the segmentation and parsing of the data.[13]

With the growing presence and salience of privacy regulation such as GDPR and the California Consumer Privacy Act, however, it seems likely that privacy regulation may allow data to become more excludable. In Miller & Tucker (2014),[14] we discuss how hospitals often limit the portability of patient data or intentionally "silo" data while citing privacy concerns as their motivation for doing so. However, we show that the

---

12 Neumann, N., C. Tucker & T. Whitfield (2019). "Frontiers: How Effective Is Third-Party Consumer Profiling? Evidence from Field Studies Nico Neumann, Catherine E. Tucker, and Timothy Whitfield Marketing Science 2019 Vol. 38 No 6, pp. 918-926.

13 https://ec.europa.eu/competition/antitrust/cases/dec_docs/38044/38044_15_5.pdf.

14 Miller, A. & Tucker C., (2014, January). Health information exchange, system size and information silos. Journal of Health Economics 33 (2), 28.

likelihood of siloing data under the guise of privacy regulation often seems to reflect the competitive structure of the local market. I modeled this as well in Campbell et al. (2015),[15] where we showed that the more strenuous the opt-in requirements are, the more likely it is that consumers will opt into sharing their data with just a few firms, rather than being willing to share data with new firms or potential entrants.

As I point out in a new paper,[16] in the past five years there appears to be evidence that in some industries where privacy concerns are particularly salient, concentration of firms increases. One industry I highlight where this has occurred is that of educational technology. The complex nature of student privacy has potentially facilitated two firms (Apple and Google) being able to supply a great deal of educational software to K-12 schools, because school administrators feel reassured that they are sufficiently privacy-compliant compared to educational technology startups in this space.

---

15 Campbell, J., Goldfarb, A. & Tucker, C. (2015), "Privacy regulation and market structure," Journal of Economics & Management Strategy 24 (1), 47.

16 See https://www.brookings.edu/wp-content/uploads/2019/12/ES-12.04.19-Marthews-Tucker.pdf.

THE IMPULSE TO CONDEMN THE STRANGE: ASSESSING BIG DATA IN ANTITRUST

BY ALEXANDER KRZEPICKI, JOSHUA D. WRIGHT & JOHN M. YUN[1]

1 J.D. Candidate, 2021, Antonin Scalia Law School, George Mason University; University of Virginia, B.S., Electrical Engineering, 2014; University Professor of Law and Executive Director, Global Antitrust Institute, Antonin Scalia Law School, George Mason University; Associate Professor of Law and Director of Economic Education, Global Antitrust Institute, Antonin Scalia Law School, George Mason University.

I. INTRODUCTION

An emerging refrain in antitrust dialog is that the accumulation and use of big data is a unique and particularly troublesome entry barrier, worthy of antitrust scrutiny. Yet, it seems that both the concept of big data and entry barriers continue to be used in a highly casual and superficial manner. Antitrust is a fact-intensive area of law, given the necessity to both understand a business practice (including its potential harms and benefits) and make forecasts of market performance. While antitrust jurisprudence has developed reasonable measures to facilitate such analyses — such as condemning price fixing as a *per se* violation — conduct such as vertical integration, resale price maintenance, and exclusive deals rightly require substantive inquiries to determine the ultimate competitive impact. Though some would argue that the courts and agencies are committing too many false positives or negatives, in the end, there is broad agreement from serious antitrust practitioners and scholars that a rule of reason analysis requires the avoidance of reflexive labels.

In this article, we argue that big data should properly be considered a two-stage process. In stage one, a firm collects the data. In stage two, a firm transforms the data into some benefit that ultimately increases profitability. This classification matters because it allows us to link big data to concepts that antitrust is already familiar with — namely, economies of scale, learning by doing, and research & development.

By linking big data with the familiar, we hope to avoid a common tendency in antitrust to condemn the strange. The history of antitrust is littered with examples where scholars, agencies, and courts have been quick to condemn new and irregular practices.[2] For instance, the following practices were subject to heavy anticompetitive judgment at one point, but are today almost universally acknowledged as having the potential to increase welfare: horizontal mergers in unconcentrated markets, vertical mergers, resale price maintenance, exclusive territories, and price discrimination.

We also discuss whether big data should be considered an entry barrier, which, in a broad and abstract sense, measures the relative difficulty of obtaining necessary inputs to production. For instance, if a firm monopolizes lithium, which is a critical and rare resource, then it can raise the price to a level that severely hinders its use. In a similar way, monopoly control over an idea due to a patent creates greater scarcity — even if only temporarily — while potentially raising prices and putting up a "barrier" to those who seek to enter. Yet, the terms "entry barriers" and "barriers to entry" are not well defined in antitrust and are used in different ways by different people.[3] Therefore, we argue, as many have previously argued, that labeling an input as an entry barrier is generally unhelpful. Rather, what is necessary is a full-fledged entry analysis — as outlined in the *Horizontal Merger Guidelines*.[4]

Ultimately, the lesson that continues to be repeated in antitrust is that the impulse to condemn the strange as anticompetitive causes tangible harm to consumers. Big data, and, increasingly, artificial intelligence and machine learning, is receiving this treatment now. Circumspection and a visit from the ghost of antitrust past counsel some degree of prudence in how we incorporate big data into antitrust law.

II. BIG DATA: A TWO STEP PROCESS

In order to understand the role of big data in production, let us begin with a concept that is familiar in antitrust. Economy of scale is the idea that the average cost of production falls when, for instance, a firm makes 1,000 units compared to 100 units. This fall in costs is due to a decrease in input costs as the firm produces more output. In a sense, it is a residual benefit of higher production levels — for instance, it allows a firm to spread fixed (including sunk) costs over significantly more output and to purchase inputs at bulk discounts, which lower the average total cost. In a similar, but slightly different, way, the *collection* of big data is a residual of consumption. For example, when a multisided platform (e.g. Uber, Lyft, Juno) brings together two groups for a mutually beneficial exchange (e.g. passengers and drivers), this creates value for the two groups as well as the platform. During this exchange, the platform can also collect data. Unlike economies of scale, however, the mere collection of big data

---

2 Professor Ronald Coase, in discussing the intersection of industrial organization and antitrust, finds the desire to be "of service to one's fellows" in the realm of public policy has created a tendency that "has discouraged a critical questioning of the data and of the worth of the analysis, leading the many able scholars in this field to tolerate standards of evidence and analysis which, I believe, they would otherwise have rejected." Further, "the association of the study of industrial organization with antitrust policy has created a disposition to search for monopolistic explanations for all business practices whose justification is not obvious to the meanest intelligence." See Ronald H. Coase, *Industrial Organization: A Proposal for Research*, *in* 3 Economic Research: Retrospect and Prospect: Policy Issues and Research Opportunities in Industrial Organization 66, 68 (Victor R. Fuchs, ed., 1972).

3 See Dennis W. Carlton, *Barriers to Entry*, 1 Issues in Competition L. & Pol'y 601 (2008).

4 U.S. Dep't of Justice & Fed. Trade Comm'n, Horizontal Merger Guidelines (2010).

does not inevitably provide a benefit that results in higher profits. Rather, collected data provides a *potential opportunity* for higher profits.[5] We can label this first stage, where the firm collects the data, as the "data-input" stage.

We can label the second stage as the "data-output" stage, where a firm takes the data and transforms it into something that creates value in the form of lower costs, improved quality, and innovative new products. This stage involves the task of combining the data with other resources and inputs, such as intellectual property, skilled labor, and capital infrastructure. Firms will have differential advantages and skills in this stage. By way of analogy, while academics all generally have access to the same scholarly journals, court cases, and, in many instances, data, the level of production will differ in terms of both quality and quantity, based on a variety of factors. Therefore, this second stage is more akin to learning by doing and research & development. In other words, this data-output stage is about innovation. The innovation could be, for example, improving a proprietary search algorithm or building a multi-dimensional profile of users on a social network for improved advertiser targeting. In either case, the firm is creating value that, in turn, increases profits.

Learning by doing is an economic concept that is like economy of scale in that it involves lower costs as output expands. The difference is that the lower costs are due to the cumulative effects of experience in production, which means a more efficient production process.[6] In economics parlance, economy of scale involves lower input costs at higher levels of output due, in no small part, to spreading fixed costs over more output; while learning by doing involves lower costs because the production function becomes more "efficient" — even holding the input costs constant. In other words, a firm improves its productivity, which, in turn, lowers its per unit costs.[7]

Importantly, learning by doing is not a process that automatically occurs. In their detailed examination of how learning by doing is actually implemented at an automaker, Professors Steven Levitt, John List, and Chad Syverson conclude that in a "more full-fledged view of learning by doing, a producer's experience gains do not so much cause efficiency enhancements themselves as they provide opportunities for management to exploit."[8] In the context of big data, economist Hal Varian states, "it can be somewhat misleading as it suggests that 'learning' is a passive activity that automatically happens as more output is produced. Nothing could be further from the truth. Learning by doing necessarily requires investment in data collection, analysis, and experimentation."[9]

In a similar way, research & development ("R&D") involves investing dedicated resources to generate new intellectual property, products, and processes. While a more successful firm will have greater capacity, experience, and accessible financial capital to invest in this process of generation, it still requires considerable risk, skill, and persistence to engage in successful R&D.

Conceptually, this need to expend resources and effort to successfully engage in learning by doing and R&D is a useful mapping for the data-output stage of big data. It suggests we should be skeptical about claims that big data somehow poses unique problems from an antitrust perspective. Like learning by doing and R&D, all firms have the capability and opportunity to use big data to improve profits through higher quality products or lower costs.

Of course, larger firms will have more data — just as larger firms will have more output and, likely, revenue. Yet, more output and revenue do not necessarily translate into anticompetitive outcomes — nor does having more data. Further, having more data confers a potential benefit with diminishing returns. The degree to which diminishing returns become a serious factor in the analysis will differ by firm and industry — yet the general principle remains.

---

5 See, e.g. Hal Varian, *Use and Abuse of Network Effects*, *in* Towards a Just Society: Joseph Stiglitz and Twenty-First Century Economics 232 (Martin Guzman ed., 2018) ("Mere data by itself doesn't confer a competitive advantage; that data has to be translated into information, knowledge, and action.").

6 For one of the earliest formulations of the concept as it applies to firms, see Kenneth J. Arrow, *The Economic Implications of Learning by Doing*, 29 Rev. Econ. Studies 155 (1962).

7 This is just one formulation of learning by doing. See Peter Thompson, *Learning by Doing*, *in* Handbook of Economics of Innovation (Bronwyn Hall & Nathan Rosenberg, eds., 2010).

8 See Steven D. Levitt, John A. List & Chad Syverson, *Toward an Understanding of Learning by Doing: Evidence from an Automobile Assembly Plant*, 121 J. Pol. Econ. 643, 647 (2013); see also John M. Dutton & Annie Thomas, *Treating Progress Functions as a Managerial Opportunity*, 9 Acad. Mgmt. Rev. 235 (1984).

9 Varian, *supra* note 5, at 229–30.

III. BIG DATA AS AN ENTRY BARRIER

The concept of entry barriers has a long history in economics and antitrust. Despite this history, there is still no universal agreement on its use.[10] The term "barriers to entry" is generally used in one of two ways. Professor Dennis Carlton summarizes this idea when he states, "Trying to use 'barriers to entry' to refer to both the factors that influence the time it takes to reach a new equilibrium and to whether there are excess long-run profits is confusing."[11] The former definition is consistent with the analysis of entry in the *Horizontal Merger Guidelines*, which involves assessing the timeliness, likelihood, and sufficiency of entry.[12] The latter definition is more in line with the economics literature on barriers.[13]

The danger in labeling a factor of production, including big data, as a "barrier to entry" is a lack of clarity regarding which definition one is considering. All business ventures involve cost — including the cost of entry. Common examples include legal and regulatory costs, licensing and developing intellectual property, expenditures on specialized equipment, and hiring skilled labor. Merely identifying a set of costs that must be incurred to achieve entry and labeling them as "entry barriers" serves no real purpose. Either the term barriers to entry is explicitly stated and the welfare consequences evaluated,[14] or, as Carlton recommends: "rather than focusing on whether an entry barrier exists according to some definition, analysts should explain how the industry will behave over the next several years . . . [which] will force them to pay attention to uncertainty and adjustment costs."[15]

Consequently, it makes little sense to label big data as a barrier to entry and thereby treat it as an inevitable impediment to competition and consumer welfare.[16] Effective investments in big data (along with machine learning and artificial intelligence) can certainly create competitive distance between rivals. Yet this distance is a byproduct of competition on the merits and, as numerous examples confirm (including the well-documented replacement of incumbents in numerous digital markets), is not necessarily an impediment to entry by innovative new firms.[17] Rather than labeling big data as a barrier to entry, the focus should be on assessing what big data helps a firm accomplish — either in a welfare-enhancing or welfare-reducing manner.

In sum, a laudable goal in antitrust is to replace designating inputs as "barriers to entry" with a more fruitful, and relevant, entry analysis. This is the approach that the U.S. antitrust agencies have adopted, where the aim is to thoroughly "examine the timeliness, likelihood, and sufficiency of the entry efforts an entrant might practically employ."[18] What must be avoided are focusing on mere possibilities — both in terms of having no entry and having easy entry. These are shortcuts that provide policymakers and courts little ultimate guidance.

While the state of entry analysis in the courts is beyond the scope of this article, it is worth noting that the lack of precision and vague notions of "entry barriers" is a problem there, too. According to Professor Daniel Lazaroff, "the Supreme Court has really never provided a comprehensive analysis of barriers to entry and their role in interpreting the Sherman, Clayton, and Federal Trade Commission Acts. Rather, the Court has periodically referenced entry barriers in antitrust cases, resulting in a somewhat cryptic and uncertain message to lower courts, litigants

---

10 See, e.g. W. Kip Viscusi, John M. Vernon & Joseph E. Harrington, Jr., Economics of Regulation and Antitrust 168 (4th ed., 2005) ("There is perhaps no subject that has created more controversy among industrial organization economists than that of barriers to entry. At one extreme, some economists argue that the only real barriers are government related . . . At the other end of the spectrum, some economists argue that almost any large expenditure necessary to start up a business is a barrier to entry.").

11 Carlton, *supra* note 3, at 606.

12 Horizontal Merger Guidelines, *supra* note 4, § 9.

13 For an overview of the evolution of economic thought on barriers to entry, see R. Preston McAfee, Hugo M. Mialon, & Michael A. Williams, *What is a Barrier to Entry?*, 94 Am. Econ. Rev. 461 (2004).

14 See Carl C. von Weizsäcker, *A Welfare Analysis of Barriers to Entry*, 11 Bell J. Econ. 399 (1980).

15 Carlton, *supra* note 3, at 615. Similarly, Demsetz observed that conditions frequently considered barriers to entry, such as scale economies, capital requirements, and advertising expenditures, are not the fundamental source of barriers; the fundamental barriers are rather the cost of information and the uncertainty that an entrant has to overcome. See Harold Demsetz, *Barriers to Entry*, 72 Am. Econ. Rev. 47 (1982).

16 Professors Anja Lambrecht & Catherine Tucker arrive at this conclusion through a slightly different approach. See Anja Lambrecht & Catherine Tucker, *Can Big Data Protect a Firm from Competition?*, CPI Antitrust Chronicle, January 2017, at 8 ("For a wide range of examples from the digital economy we demonstrate that when firms have access to big data, at least one, and often more, of the four criteria which are required for a resource to constitute a sustainable competitive advantage are not met.").

17 See Competition Bureau Canada, Big Data and Innovation: Implications for Competition Policy in Canada 14 (2017), https://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/vwapj/CB-Report-BigData-Eng.pdf/$file/CB-Report-BigData-Eng.pdf ("Developing valuable data through competition on the merits does not run afoul of the Act even if it results in significant market power. For example, a firm can create market power by developing a high-quality product or an efficient production process.").

18 Horizontal Merger Guidelines, *supra* note 4, § 9.

and students of antitrust law."[19] A brief survey of recent antitrust cases confirms Professor Lazaroff's observation that the current treatment of entry barriers remains relatively perfunctory and lacking in clarity. For instance, conclusions about what constitutes an entry barrier verge on the contradictory.[20] Some courts have characterized the existence of economies of scale in a market as an indication of barriers to entry; although, it would not be considered a barrier to entry under influential definitions, such as from Professor George Stigler.[21] Even in cases where the discussion of entry barriers is more thorough, the lack of an accepted economic framework leaves courts fumbling in the dark when it comes to deciding the question.[22] All told, this is an area of antitrust law ripe and yearning for a more formalized and rigorous scheme of analysis, grounded in sound economics.

IV. CONCLUSION

As interest in antitrust policy is expanding, the commitment to uphold the integrity of economic and legal analyses should be renewed. Little is to be gained from moving antitrust from a fact-based area of law to a more reflexive, rhetorical area. Consistent with the prior literature on big data, we argue that big data should be considered a two-stage process in antitrust analysis. This consideration allows antitrust law to properly consider big data in the context of a larger entry analysis.

---

19 Daniel E. Lazaroff, *Entry Barriers and Contemporary Antitrust Litigation*, 7 U.C. Davis Bus. L.J. 1 (2006), https://blj.ucdavis.edu/archives/vol-7-no-1/Entry-Barriers-and-Contemporary-Antitrust-Litigation.html.

20 Compare *GDHI Mktg. LLC v. Antsel Mktg. LLC*, No. 18-CV-2672-MSK-NRN, 2019 WL 4572853 at *9 n.5 (D. Colo. Sept. 20, 2019) ("The mere cost of capital is not a barrier to entry.") with *Philadelphia Taxi Ass'n, Inc v. Uber Techs., Inc.*, 886 F.3d 332, 342 (3d Cir. 2018) ("Entry barriers include . . . high capital costs.").

21 See, e.g. *N.M. Oncology v. Presbyterian Healthcare Servs.*, No. CV 12-00526 MV/GBW, 2019 WL 6040036, at *7 (D.N.M. Nov. 14, 2019) ("[T]here are significant and continuing barriers to entry into the relevant markets. First, there can be no dispute that any new insurer would need to build a provider network."); see George J. Stigler, The Organization of Industry 67 (New ed. 1983); John M. Yun, *Antitrust After Big Data*, 4 Criterion J. Innovation 407, 421 (2019).

22 See *Buccaneer Energy (USA) Inc. v. Gunnison Energy Corp.*, 846 F.3d 1297, 1316–17 (10th Cir. 2017).

ESSENTIAL FACILITIES FALLACY:
BIG TECH, WINNER-TAKE-ALL MARKETS, AND
ANTICOMPETITIVE EFFECTS

BY JOHN PECMAN, PAUL A. JOHNSON & JUSTINE REISLER[1]

1 John Pecman is a Senior Business Advisor in the Antitrust/Competition & Marketing group at Fasken. Paul A. Johnson is Owner at Rideau Economics. Justine Reisler is an associate in Fasken's Antitrust/Competition & Marketing Group. The opinions expressed in this article are our own.

I. INTRODUCTION

The Big Tech superstar firms, which is a term used by many to refer to Google, Amazon, Facebook, Apple, and sometimes Microsoft, have grown to be among the largest companies in the world principally through innovation, data harvesting and the popularity of their products. Over the past few years, there have been calls for action by populists, the press and politicians to reduce the power of the so-called "dataopolies."[2]

      The business models of Google and Facebook, for example, operate through multi-sided platforms. They offer products to consumers on one side of the platform for free and receive the consumers' attention and data, which they monetize on the other side of the platform by selling targeted advertising. These business models attract a large number of users which enables the companies to build banks of data. Data in this business model is an important input for value creation. Thus, as with anything that is used to create value, data can constitute a barrier to entry if its creation, acquisition, or maintenance is costly.

      A firm that uses data may benefit from a "user feedback loop" to outcompete rivals. User feedback loops occur when firms use data collected from their users to improve the quality of their product, in turn attracting more users, which leads to further quality improvements and yet more users. This user feedback loop can tip markets in favor of winner-take-all outcomes that may appear stable. In other words, there may be a danger that markets do not self-correct through the competitive process as in "natural monopolies."

      This view, that the Big Tech superstar firms have durable and significant market power, has been articulated by four recently published reports: the European Union's ("EU") *Competition for the digital era*, final report,[3] the United Kingdom's ("UK") *Unlocking Digital Competition: Report of the Digital Competition Expert Panel* (the "Furman Report"),[4] Australia's *Digital platforms inquiry*, final report[5] and the University of Chicago's *Stigler Committee on Digital Platforms*, final report (the "Stigler Report").[6] The Furman report called for a market study, which is currently being carried out, into the UK's digital advertising market encompassing the entire value chain, using its investigatory powers to examine whether competition is working effectively.[7]

      This is the context in which calls have been made to treat data, more or less, as an essential facility.[8] That is, there have been calls to impose data sharing regulations on companies that have, use, or collect data to prevent them from exercising market power and to reduce barriers to entry. For instance, the UK's Competition and Markets Authority ("CMA") in its interim report on online platforms and digital advertising, considers a number of interventions to regulate the activities of online platforms funded by digital advertising, such as "a requirement that Google provide click-and-query data to rival search engines."[9] The CMA has taken the view that a lack of comparable scale in click-and-query data is likely to be a key factor that limits the ability of other search engines to compete with Google. The CMA proposal would essentially make click-and-query data an essential facility and impose a duty to deal on it.

---

2 Franklin Foer, *World Without Mind: The Existential Threat of Big Tech* (New York: Penguin Press, 2017); Lina M. Khan, "Amazon's Antitrust Paradox," *The Yale Law Journal* 126:3 (January 2017), available at https://www.yalelawjournal.org/note/amazons-antitrust-paradox; Nick Srnicek, "The only way to rein in big tech is to treat them as a public service," *The Guardian* (April 23, 2019) , available at https://www.theguardian.com/commentisfree/2019/apr/23/big-tech-google-facebook-unions-public-ownership; David McCabe, "Lawmakers Urge Aggressive Action from Regulators on Big Tech," *The New York Times* (September 17, 2019), available at https://www.nytimes.com/2019/09/17/technology/senate-antitrust-tech-hearing.html; Elizabeth Warren, "Here's how we can break up Big Tech," *Medium* (March 8, 2019), available at https://medium.com/@teamwarren/heres-how-we-can-break-up-big-tech-9ad9e0da324c.

3 Jacques Crémer, Yves-Alexandre de Montjoye & Heike Schweitzer, *Competition Policy for the digital era*, Final Report, European Commission, Directorate-General for Competition (2019), available at https://ec.europa.eu/competition/publications/reports/kd0419345enn.pdf at pgs. 70-72.

4 United Kingdom, *Unlocking Digital Competition: Report of the Digital Competition Expert Panel,* also known as the 'Furman Report' (March 2019), available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf at paras. 1.112 - 1.117, 2.109, 2.116 & 3.26.

5 Australian Competition & Consumer Commission, *Digital platforms inquiry*, final report (July 26, 2019), available at https://www.accc.gov.au/focus-areas/inquiries-ongoing/digital-platforms-inquiry/final-report-executive-summary: pgs. 7-10 and 57-105.

6 University of Chicago, Stigler Center, *Stigler Committee on Digital Platforms,* Final Report (September 2019), available at https://research.chicagobooth.edu/stigler/events/single-events/antitrust-competition-conference/digital-platforms-committee at pgs. 78, 119-121.

7 United Kingdom, Competition & Markets Authority, *Online platforms and digital advertising,* Market study interim report (December 2019): at para. 1.112, available at https://assets.publishing.service.gov.uk/media/5dfa0580ed915d0933009761/Interim_report.pdf.

8 Zachary Abramson, "Essential Data" *The Yale Law Journal* 124:867 (2014); Nick Srnicek, "The only way to rein in big tech is to treat them as a public service," *The Guardian* (April 23, 2019) , available at https://www.theguardian.com/commentisfree/2019/apr/23/big-tech-google-facebook-unions-public-ownership; Inge Graef, "Rethinking the Essential Facilities Doctrine for the EU Digital Economy," *Revue Juridique Themis* 53:1 (2019).

9 United Kingdom, Competition & Markets Authority, *Online platforms and digital advertising,* Market study interim report (December 2019) at para. 79, available at https://assets.publishing.service.gov.uk/media/5dfa0580ed915d0933009761/Interim_report.pdf.

The imposition of a duty to deal on an essential facility prevents the *exercise* of market power. Such a prohibition on exercising market power is distinct from a prohibition on *creating, protecting, or enhancing* market power by anticompetitive acts, which is the object of traditional antitrust enforcement. One reason that antitrust law is not normally designed to regulate the exercise of market power is because it would require competition agencies to become sector regulators, a function they are not equipped to perform. Another reason is that mandating a duty to deal can chill incentives to innovate. But whatever the intervention considered, it would seem to be uncontroversial to insist on careful consideration of how the intervention may affect consumers and competition more broadly. This consideration is particularly important with Big Tech companies that have popular products with a significant presence in the economy and in people's lives.

II. IT IS IMPORTANT TO CONSIDER DYNAMIC COMPETITION BEFORE MANDATING A DUTY TO DEAL

Contrary to some of the findings in certain recently published reports,[10] there are reasons to question whether industries where data is important are natural monopolies or that barriers to entry are durable beyond the short term. Generic calls to treat companies with large-scale data sets as natural monopolies or blanket claims that "data is the new oil" coarsely lump all data together. However, data plays very different roles in different markets. For example, data can be collected from retail commerce (e.g. Walmart has had one of the largest databases in the world[11]), from search engines, or from the "Internet of Things." Data may also be bought and sold through data brokers[12] or kept as proprietary information to be sold to clients for a price. Data is also "non-rivalrous."[13] In other words, the use of raw data by one firm does not preclude its use by other firms.

As the economist Joseph Schumpeter argued, it is difficult to predict future developments of companies and technologies. Schumpeter's famous term "creative destruction" was used to describe the process through which firms innovate to capture consumers and market share, only to be eventually displaced themselves. He posited a concern that this dynamic aspect of competition is too often overlooked in order to assess competition simplistically in a static fashion.[14] It is easy to find evidence that such a concern is still relevant today. For example, a 2007 article in *The Guardian* suggested that the social media company, Myspace, was a natural monopoly.[15] Extensive network effects were the basis for the claim that Myspace's dominant market position was unassailable similar to those being made about Facebook and Google today.

Myspace, founded in 2003, quickly saw a rapid rise in its market position. The website was a social network with individual profiles, creating networks of friends and opportunities to embed or connect to music. By May 2009, Myspace had been overtaken by Facebook. With a more user-friendly interface, Facebook's superior advertising allowed for greater innovation propelling them to the top and displacing MySpace in the social media market.[16] Apart from the advent of new superior network usages or technologies, any durable barriers to entry of the dominant digital platforms could be also be challenged on account of lower fixed entry costs associated with cloud computing, growing overall venture capital investments in the digital economy, and improved data portability such as those that may result from voluntary initiatives by Big Tech companies such as the Data Transfer Project.[17]

---

10 Jacques Crémer, Yves-Alexandre de Montjoye & Heike Schweitzer, *Competition Policy for the digital era*, Final Report, European Commission, Directorate-General for Competition (2019) at pg. 112, available at https://ec.europa.eu/competition/publications/reports/kd0419345enn.pdf; United Kingdom, *Unlocking Digital Competition: Report of the Digital Competition Expert Panel,* also known as the 'Furman Report' (March 2019) at paras. 1.65-1.92, available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf; Australian Competition & Consumer Commission, *Digital platforms inquiry*, final report (July 26, 2019) pg. 58, available at https://www.accc.gov.au/focus-areas/inquiries-ongoing/digital-platforms-inquiry/final-report-executive-summary; University of Chicago, Stigler Center, *Stigler Committee on Digital Platforms,* Final Report (September 2019) at pg. 43, available at https://research.chicagobooth.edu/stigler/events/single-events/antitrust-competition-conference/digital-platforms-committee.

11 Charles Fishman, *The Wal-Mart Effect: How the World's Most Powerful Company Really Works - and How it's Transforming the American Economy* (New York: Penguin Books, 2006).

12 Federal Trade Commission, *Data Brokers: A call for transparency and accountability* (May 2014), available at https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

13 D. Daniel Sokol & Roisin Comerford, "Antitrust and Regulating Big Data," *George Mason Law Review* 23:5 (Fall 2016) pg. 1137.

14 Joseph A. Schumpeter, *Can Capitalism Survive? Creative Destruction and the Global Economy* (New York: Harper Perennial, 2009). Originally published as *Capitalism, Socialism, and Democracy.*

15 Victor Keegan, "Will MySpace Ever Lose its Monopoly?" *The Guardian* (February 8, 2007).

16 Ryan Bourne, "Is This Time Different? Schumpeter, the Tech Giants, and Monopoly Fatalism," CATO Institute (June 17, 2019), available at https://www.cato.org/publications/policy-analysis/time-different-schumpeter-tech-giants-monopoly-fatalism#null.

17 Data Transfer Project, available at https://datatransferproject.dev.

Notwithstanding such potential disruption, economies of scale and scope do appear to be particularly strong in relation to the accumulation and use of certain types of data relating to consumer behavior. A potential new entrant in a given market may face significant barriers to entry and expansion from incumbents who use large-scale datasets to improve the quality of their own products and services. Big Tech companies can apply machine learning to extensive data sets to improve their offerings and expand their activities into new areas. Thus, incumbent firms may be able to use machine learning trained on large datasets to cement their position by attracting more users by improving their services; they may also end up making more money by better targeting their advertising. By limiting a rival's access to such data, dominant firms can provide strong incentives for consumers to use their services exclusively.

III. THE EXISTING ANTITRUST FRAMEWORK MAY LEAD TO BETTER OUTCOMES THAN REGULATION

The traditional antitrust framework has nothing to do with ensuring that new entrants have the right to succeed or that firms take (subjective) steps that support democracy, fairness or freedom. That is not to say that such concerns are not valid; it is just that they have traditionally been more of a political issue than a technocratic antitrust issue. For instance, there may be a public interest in creating legislation or regulations that protect privacy, media diversity and democratic institutions, but these are not antitrust issues. What is interesting is that certain solutions to antitrust problems, such as giving consumers more autonomy over their data through portability, may be a solution to both an antitrust and a public interest concern. There is nothing wrong with killing two birds with one stone, but it is also important to be precise about the problems that each solution is intended to solve. Competition is best promoted with minimal regulation and in keeping with the rationale for economic regulation, which is rooted in addressing market failures such as those arising from natural monopolies or externalities.

The benefit to using the traditional antitrust framework is that it avoids micro-managing the technology sector, a complex and rapidly evolving sector of the economy. Of course, there are large sectors of the economy that are currently regulated (e.g. a municipal electricity distribution company may have its prices, the services it offers, and investments it makes decided by regulation). But traditionally regulation is only imposed in industries where the cost of entry is so great, or the duplication of facilities would be so wasteful, that some degree of monopoly is considered unavoidable.[18] For instance, public utilities are regulated based on their significant economies of scale such that one firm can provide service to a market at a lower cost than two or more competing firms. It is also important to note that most regulation is imposed on industries where the nature of goods or services is unchanging (e.g garbage pick-up, electricity distribution) relative to a rapidly evolving technology sector.

While we believe that calls for regulation to fix perceived problems in big tech are drastic and should be questioned, we believe that it is appropriate to examine the existing antitrust framework in light of current concerns. In other words, it is reasonable to ask whether the existing antitrust framework could potentially benefit from being "fine-tuned." Such fine tuning has happened in the past and it is natural to expect it will continue in the future. Past examples include (a) moving from *per se* condemnation of certain vertical restraints (e.g. resale price maintenance) in favor of a rule of reason approach in the United States,[19] (b) the European Commission's adoption of a new substantive test for merger control, the so-called SIEC ("significant impediment of effective competition") test[20] and (c) numerous revisions to Canada's competition legislation in 1986 including the introduction of civil reviewable practices to address mergers, abuse of dominance, refusal to deal, exclusive dealing, and tied selling, removing most competition matters from the cumbersome criminal regime.[21]

This article concerns a specific kind of "fine tuning" that might be useful to consider in light of the concerns about data usage by Big Tech. It examines the legal test and standard of proof on which the anticompetitive effects of a transaction or conduct are proven and met, and by whom.

---

18 Walter Adams, "The Role of Competition in the Regulated Industries," *Journal of Reprints for Antitrust Law and Economics* 9.2 (1978), reprinted from 48 American Economic Review 527 (1958) pg. 457.

19 *Leegin Creative Leather Products, Inc. v. PSKS Inc.,* 551 U.S. 877 (2007); Iacobucci, E. & Winter, R.A., "Vertical Restraints Across Jurisdictions," *The Oxford Handbook of International Antitrust Economics*, Roger D. Blair & D. Daniel Sokol, Eds' (Oxford: Oxford University Press, 2014) pgs. 351-386.

20 Council Regulation (EC) No 139/2004 of 20 January 2004, [2004] OJ L24, 1; "Commission Notice - Guidelines on the Assessment of Horizontal Mergers under the Council Regulation on Control of Concentrations between Undertakings," DG COMP, 28 January 2004; Lars-Hendrik Roller & Miguel De La Mano, "The Impact of the New Substantive Test in European Merger Control," *European Competition Journal* 2:1 (April 2006).

21 Bill C-91, An Act to establish the Competition Tribunal and to amend the Combines Investigation Act and the Bank Act and other Acts in consequences thereof, 1st Sess, 33rd Parl, 1985 (assented to June 17, 1986), SC 1986, c 26; Sandra Walker & Michael Kirby, "Institutional Overview and Statutory History," *Fundamentals of Canadian Competition Law,* 3rd ed, James B Musgrove ed. (Toronto: Carswell, 2015): pgs. 24-26.

IV. ASSESSING EFFECTS IN BIG TECH

Much antitrust enforcement requires an assessment of effects. For example, in Canada there is no injunctive relief from an alleged anticompetitive act without showing anticompetitive effects. There is also no enjoining of a merger unless it can be shown that the merger is likely to cause a substantial lessening or prevention of competition ("SLPC"). The SLPC test is very similar to the 'substantial lessening of competition' ("SLC") test, which is used by most other common law jurisdictions (*e.g.* the United States, the UK, Ireland, Australia, and Singapore).

That is not to say that showing effects is critical to all antitrust enforcement. For example, in most jurisdictions, cartel conduct is a *per se* offence and condemned without proof of anticompetitive effects. And in Europe, even some types of agreements between non-competitors are treated as "by object" infringements of competition law that do not require a showing of effects.[22] Nevertheless, a significant amount of antitrust enforcement, in most jurisdictions, requires a showing of anticompetitive effects.

In light of the importance of demonstrating anticompetitive effects, the remainder of this article considers the current legal test required to show anticompetitive effects and suggests that, in situations where economies of scale, scope, and/or network effects are important — as some have claimed about certain tech markets — that test may be overly onerous for antitrust plaintiffs when markets are characterized by a "winner-take-all" dynamic.

A. Winner-take-all Markets

Some markets support the existence of multiple, even a great many, competitors. For example, driving on a highway in Canada reveals that a great many companies offer trucking services. Other markets are different in that a substantially smaller number of producers typically exist. For example, as of October 2019, Android and iOS accounted for nearly all the share of a "worldwide mobile operating system market" with Android accounting for about three times iOS's share (77 percent and 22 percent).[23] That was not always the case, however, as Blackberry's share peaked around 20 percent in 2010 although by early 2017 its share was 0.0 percent.[24]

In such markets, which we will call "winner-take-all markets," there is competition *for* the market. By contrast, in traditional markets, there is competition *in* the market.[25] Some have claimed that competition for the market — as opposed for competition in the market — is a key feature of markets in which big tech firms are active.[26]

Winner-take-all markets support a small number of competitors because large firms can create much more value than smaller firms. Consider an extreme example: the benefit of a water and sewer authority being "big" is so great that we resign ourselves to regulating it by specifying the nature of what it sells and at what price. Any of the following three types of efficiencies may be why bigger firms deliver greater value than smaller firms.

- *Economies of scale* are present when firms that sell more of a particular product incur costs that are lower in proportion to firms that sell less of a particular product. The existing water and sewer authority can deliver service to a new house for substantially lower costs than could a new entrant who would not only have to connect the house to a network of supply and sewer lines, but also construct that network and construct water filtration and wastewater collection and treatment facilities.

---

22 European Commission, "Guidance on restrictions of competition 'by object' for the purpose of defining which agreements may benefit from the De Minimis Notice" Commission Staff Working Document, Brussels (June 25, 2014), available at https://ec.europa.eu/competition/antitrust/legislation/de_minimis_notice_annex.pdf.

23 "Mobile Operating System Market Share Worldwide: Nov 2018 - Nov 2019," *Statcounter, Global Stats* (December 2019), available at https://gs.statcounter.com/os-market-share/mobile/worldwide.

24 Jake Swearingen, "BlackBerry Now Controls 0.0 Percent of the Smartphone Market," *New York* (Feb. 15, 2017), available at http://nymag.com/intelligencer/2017/02/blackberrys-global-market-share-is-now-0-0.html.

25 Evans, D.S., & R. Schmalensee, "Some Economic Aspects of Antitrust Analysis in Dynamically Competitive Industries," *Innovation Policy and the Economy* 2 (2002) 1-49; Paul Andrew Geroski, "Competition in Markets and Competition for Markets," *Journal of Industry, Competition and Trade* 3 (2003): 151-166; Joshua S. Gans, "When is Static Analysis a Sufficient Proxy for Dynamic Considerations? Reconsidering Antitrust and Innovation," *Innovation Policy and the Economy* 11 (2011): 55-78.

26 Charles McConnell, "The digital economy means competition for the market, says Scott Morton," *Global Competition Review* (November 13, 2017), available at https://globalcompetitionreview.com/article/usa/1150021/the-digital-economy-means-competition-for-the-market-says-scott-morton.

- *Economies of scope* are present when firms that sell more products incur costs that are lower in proportion to firms that sell fewer products. Some automobile manufacturers take advantage of economies of scope when they produce multiple automobile models based on common engineering designs. For example, Volkswagen's "Modular Transverse Matrix standardizes many vehicle component parameters across brands and vehicle classes. It essentially means that one platform can be used to build a wide variety of cars. This means the highest levels of technology can be shared across the Volkswagen Group, and production costs can be reduced while production flexibility is increased."[27]

- *Network effects* are present when consumers value a particular product more as more consumers also consume that particular product. In contrast to economies of scale or scope, which are characteristics of supply, network effects are a characteristic of demand. Perhaps the most obvious example is a social media product like Facebook where users benefit when more users join the platform.[28]

To exploit these efficiencies, a certain amount of scale is necessary. Thus, in winner-take-all markets — that is in markets where some or all of these efficiencies are particularly strong — a successful firm must be a big firm and, hence, its presence or absence will have very substantial competitive effects. In other words, winning competition *for* a market is likely to result in very substantial competitive effects. Those effects are likely to be much less substantial if competition is only *in* the market. Consider, for example, the likely competitive effects of a merger between iOS and Android compared with the likely effects of a merger of firms that offer trucking services in Canada. Clearly, the former would be more substantial - measured in any meaningful way - than the latter.

B. Assessing Anticompetitive Effects in Winner-take-all Markets

Analysis of anticompetitive effects under an SLC test can operate very differently in winner-take-all markets compared with traditional markets. In short, while an SLC test can work well in traditional markets — in the sense that some conduct or acquisition will be enjoined only if welfare is harmed *in expectation* — an SLC test may not work so well in winner-take-all markets. The reason is that effects will usually be small in scale in traditional markets so that the likelihood of those effects is a reasonable proxy for how they would affect welfare *in expectation*. By contrast, effects will usually be of enormous scale in winner-take-all markets so dismissing such effects whenever they are deemed unlikely to occur will lead to a failure to enjoin a merger or conduct that lowers welfare substantially *in expectation*.

The words "in expectation" above are used in the mathematical sense of "expected value" and are important to appreciate. The expected value of an event is defined to be the product of the event's probability and the scale of its effect, if it occurs. Symbolically, if an event occurs with probability p and the effect of that event, if it occurs, is denoted x, the expected value of the event is equal to px. To illustrate, suppose an anticompetitive effect of $10,000 occurs with probability 10 percent but, otherwise (i.e. with probability 90 percent) that effect is zero. In that case, the expected value is equal to $10\% \times \$10,000 + 90\% \times \$0 = \$1,000$.

A fundamental property of an expected value is that a reduction in probability can be perfectly offset by an increase in the scale of the effect. While perhaps abstract when stated in this way, this property is key to what follows. To illustrate, the following statements all imply the same expected value of anticompetitive effects: 1) an anticompetitive effect of $10,000 occurs with probability 10 percent and $0 otherwise, 2) an anticompetitive effect of $100,000 occurs with probability 1 percent and $0 otherwise, and 3) an anticompetitive effect of $1,000 occurs with probability 100 percent.

The fungibility between probability and scale of effect is universally appreciated by humans and businesses in the ordinary course. Examples are numerous. Many people wear their seat belt when driving even though the probability of an accident is very small; they realize that the costs of an accident without wearing a seat belt can be very large. A college applicant may apply to "sure thing" colleges as well as "stretch" colleges. While admission at a very prestigious university may be unlikely, or a "stretch," the benefits of gaining acceptance may be perceived to be very large and offset the lower likelihood of acceptance. Insurance companies charge positive premiums to those they insure even though accidents are unlikely. Conversely, those that are insured pay those premiums even though they are unlikely to need the coverage. Pharmaceutical companies will pursue expensive and lengthy research and development efforts on a particular molecule despite the fact that those efforts may be extremely unlikely to result in commercial success.

---

27 "Details on the Volkswagen Group's Modular Transverse Matrix (MQB)" *Volkswagen UK media site* (2018), available at https://vwpress.co.uk/technology/technology-volkswagen-groups-mqb.

28 Such an example is a direct network effect. Other products exhibit what are called indirect network effects. For example, an Amazon Marketplace consumer derives no direct benefit as other consumers use Amazon Marketplace, but that consumer does derive benefit from an increase in the number of merchants that sell on Amazon Marketplace. And importantly, on Amazon Marketplace, the number of merchants may grow with the number of consumers. Thus, an Amazon Marketplace consumer derives an indirect benefit from an increase in the number of Amazon Marketplace consumers. Indirect network effects usually operate across different sides of a platform; direct network effects operate within a single set of users, for example, on a given side of a platform.

An SLC test operates very differently from an expected value because if effects are not found to be likely, they are disregarded entirely. In other words, there is no fungibility between the likelihood of an event and the scale of its effects. To illustrate, in only the third case described above (an anticompetitive effect of $1,000 occurs with probability 100 percent) was anticompetitive harm likely.

To see the implications further, consider how an SLC test functions in a typical traditional market and a typical winner-take-all market, on a balance of probabilities:

- Case 1: entry in a traditional market (where there is competition *in* the market). Suppose that an entrant is poised to enter a market and the procompetitive effect of its entry is uncertain but equally likely to be anywhere between $0 and $100. (To visualize, suppose that the numbers $0, $1, $2…$100 are written on 101 balls that are placed in an urn from which a single ball is drawn at random. The number on the ball that is drawn represents the procompetitive effects of entry that are realized.) This case reflects competition in the market because the entrant's effect on the market can be small. Moreover, the entrant is almost certain to make at least some sales. Thus, we are almost certain that entry will cause at least some procompetitive benefit: there is a 99 percent probability that the procompetitive effects of entry will be at least $1. Thus, on a balance of probabilities (i.e. a probability greater than 50 percent), entry will have some procompetitive effects. The expected value of the procompetitive effects of entry in this case is $50.[29]

- Case 2: entry in a winner-take-all market (where there is competition *for* the market). Suppose that an entrant is poised to enter a market and the procompetitive effects of its entry is equal to $10,000 with a probability of 10 percent and $0 with probability of 90 percent. This setting reflects competition for the market because the entrant either fails entirely or has a relatively large effect. This case supports the following statement: "On a balance of probabilities (i.e. a probability of greater than 50 percent), the procompetitive effect of entry will be $0." Nevertheless, the expected value of this entry in this case ($1,000) is substantially greater than in Case 1 ($50). Interestingly, the statement "On a balance of probabilities, the procompetitive effect of entry will be $0" remains true *regardless of the scale of the effects of entry*. For example, if the effects of entry were increased 10-fold, the expected value of entry would increase 10-fold but, on a balance of probabilities, the procompetitive effect remains zero.

Now consider an anticompetitive act by an incumbent that somehow precludes any successful entry. In other words, any attempted entry results in no procompetitive effects. This anticompetitive act would cause harm on a balance of probabilities for Case 1, but not for Case 2 even though precluded entry in Case 2 has substantially larger effects, in expectation, than in Case 1. Moreover, that remains true even as the scale of the effect in Case 2 becomes arbitrarily large. Thus, a competition enforcement agency that is required to show the existence of anticompetitive effects from precluding entry on a balance of probabilities could only successfully challenge the conduct in the case of competition in the market (i.e. Case 1) and not in the case of competition for the market (i.e. Case 2).

The discussion above highlights the fundamental point of this section and bears repeating. An enforcement rule that ignores an event regardless of the scale of its effect so long as it is deemed unlikely is problematic. Such an enforcement rule would accept *any* loss in welfare so long as it is deemed unlikely to occur (i.e. will occur with a probability of less than 50 percent). Humans and businesses certainly do not use these types of rules to make decisions. No (sane) human would accept a small payment in exchange for engaging in behavior that causes severe injury with a small probability (e.g. Russian roulette).

We suspect that this point is relevant to the current debate about whether competition law is up to the task of preventing harmful acquisitions or conduct by big tech companies. In the next section we discuss several propositions to change competition law enforcement and assess how they would operate in light of these observations.

V. PROPOSITIONS TO CHANGE COMPETITION LAW IN RESPONSE TO BIG TECH

As detailed above, a substantial share of the current debate about antitrust and big tech concerns the use of regulation. While regulation is certainly necessary to address certain non-competition problems, we are dubious about calls to regulate big tech companies just as a utility. A more useful discussion is to consider whether current antitrust enforcement can be modified in ways that address specific competition concerns in these markets in light of our observations above.

---

29 To see this without resorting to calculus, suppose that effects take on the integer values $0, $1, $2… $100 with equal probability. This implies that each integer occurs with probability (1/101). Summing the product of the values and probabilities leads to the expected value of $50.

We perceive there to be three broad categories of calls to modify the competition law framework for proving anticompetitive effects in the context of big tech:

- Those that adopt an explicitly probabilistic approach and reject a traditional analysis of effects based on a balance of probabilities,

- Those that adopt a bifurcated approach, whereby the burden of proof or applicable legal test is applied as a function of some observable initial condition, and

- Those that uniformly lessen the burden on an antitrust plaintiff.

The probabilistic approach is exemplified by the "balance of harms" approach articulated in the Furman Report submitted to the CMA.[30] This "more economic assessment" considers both the likelihood and the magnitude of the impact of a merger or conduct allowing a competition authority "to intervene where it expects, on average, for the harm of the merger to be substantially greater than the benefits."[31] In the language introduced above, this approach proposes making the notion of expected value central to an assessment of effects. While a departure for competition law enforcement, the report points out that such an approach is already used in other areas of government: "The balance of harms test would have similarities with the government's recognized approach for making regulatory decisions, which draws on the principles of cost-benefit analysis. This can combine qualitative and quantitative analysis and judgements, with various techniques for addressing the challenges of uncertainty. This approach is frequently used for significant and complex government decisions, for example for public health proposals, environmental protection, or major infrastructure investment."[32]

Ultimately, this suggestion was rejected by the CMA due to "practical challenges in applying this kind of test in a transparent and robust way."[33] One practical challenge is estimating with sufficient precision the probabilities required in the approach.[34] For example, suppose a merger would create $10 million of benefit with certainty, but creates $100 million of harm with some probability. Thus, whether that merger's net effect, in expectation, is positive or negative depends on whether that probability is greater than or less than 10 percent. Estimating probabilities with such precision may very well be challenging. For example, it may require understanding when evidence is consistent with a 9 percent probability and when evidence is consistent with an 11 percent probability.

A bifurcated approach is illustrated by Senator Amy Klobuchar's bill, the Consolidation Prevention and Competition Promotion Act of 2019.[35] Among other things, it proposes to shift the burden of proof onto the merging parties to demonstrate that the merger does not harm competition whenever the proposed acquisition either significantly increases consolidation or involves a large transaction price or large companies.[36] (It would also change the legal test in the Clayton Act[37] for prohibiting mergers from an SLC test to a "material lessening of competition" test.)

Such a proposal is bifurcated because it involves two distinct steps. The first step involves a determination of which party bears the burden of proof and the second step involves one or the other party meeting their burden of proof on the new legal test - whether the acquisition will tend to materially lessen competition or tend to create a monopoly or a monopsony, on a balance of probabilities.

---

30 United Kingdom, *Unlocking Digital Competition: Report of the Digital Competition Expert Panel,* also known as the 'Furman Report' (March 2019), available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf. See Recommendation s 9 and 10 at pgs. 13, 98-101.

31 *Ibid*, at para 3.96. Note that the report only suggested adopting the balance of harms test for merger review.

32 *Ibid*, at para. 3.91.

33 Letter from Dr. Andrea Coscelli (Chief Executive, Competition and Markets Authority) to Alex Chisholm (Permanent Secretary, Business, Energy and Industrial Strategy, Department for Business, Energy and Industrial Strategy) and Charles Roxburgh (Second Permanent Secretary, Her Majesty's Treasury), "Digital Competition Expert Panel Recommendations - CMA view" (March 21, 2019): pg. 5, available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/788480/CMA_letter_to_BEIS_-_DCEP_report_and_recommendations__Redacted.pdf.

34 Jeffrey M. Wilder, "Acting Deputy Assistant Attorney General Jeffrey M. Wilder Delivers Remarks at the Hal White Antitrust Conference," Speech, Washington, DC, (June 10, 2019), available at https://www.justice.gov/opa/speech/acting-deputy-assistant-attorney-general-jeffrey-m-wilder-delivers-remarks-hal-white.

35 S. 307 - 116th Congress (2019-2020). Available online at https://www.congress.gov/bill/116th-congress/senate-bill/307?q=%7B%22search%22%3A%5B%22S.+192%22%5D%7D.

36 The financial thresholds that define such acquisitions would be adjusted annually. As of 2019, they would implicate acquisitions greater than $5 billion in value or involving a party with assets, net annual sales, or a market capitalization greater than $100 billion.

37 Clayton Antitrust Act of 1914 (15 U.S.C. §§ 12-27).

The first step involves determining who has the burden of proof based on whether the financial threshold is exceeded. If the threshold *is* exceeded, then the merging parties have the burden of proving the acquisition will not "materially lessen competition," on a balance of probabilities. If the threshold *is not* exceeded, then it is the antitrust plaintiff's responsibility to bring sufficient evidence to prove a material lessening of competition in order to block the merger. The reversal of the burden of proof for acquisitions exceeding the threshold would be a significant change because the starting point for the court in such cases would be that the merger harms competition. This first step, determining who has the burden of proof, is unlikely to be a significant issue in a trial as it is highly factual.

The second step would either require the antitrust plaintiff to bring evidence that a merger harms competition or the merging parties to bring evidence that the merger does not harm competition, depending on whether the threshold is exceeded. Practically, both parties will bring this evidence regardless of who has the burden of proof. However, it is easier for the antitrust plaintiff to win its case when the court is starting with the presumption that a merger harms competition and the merging parties have to show it does not create a material lessening of competition. And on the other hand, it is easier for the merging parties to win their case when the court's starting point is that the transaction does not harm competition and the antitrust plaintiff has to show that the transaction has created a material lessening of competition.

An approach that uniformly lessens the burden on an antitrust plaintiff is one where a lower legal standard applies. For example, some have proposed adopting a strict structuralist approach to antitrust, which, for example, makes illegal *any* merger that creates a firm with 10 percent or more market share.[38] We have already mentioned Amy Klobuchar's proposal to change the legal test for prohibiting mergers in the United States from a "*substantial* lessening of competition" to a "*material* lessening of competition." Other suggestions to lessen the burden on antitrust plaintiffs include modifying antitrust law to relax the proof requirements imposed upon them in appropriate cases, such as those involving technology platforms. It has been proposed that courts should be more willing to permit plaintiffs to prove harm to competition by circumstantial evidence, especially where the propositions in question are not observable and there thus cannot be direct evidence.[39] It has also been proposed that courts should not presume efficiencies from vertical transactions and that crediting of efficiencies should require strong supporting evidence showing merger-specificity and verifiability.[40]

One objection to a bifurcated approach that is conditioned on some measure of size or a uniform lowering of burdens on antitrust plaintiffs is that both proposals could prevent procompetitive and efficiency-enhancing transactions. For example, the Furman Report noted the risks inherent in adopting a presumption that acquisitions by large digital companies are anticompetitive: "the majority of acquisitions by large digital companies are likely to be either benign or beneficial for consumers, though a minority may not be. Being acquired is also an important exit strategy for technology start-ups, providing significant incentives for investors to provide funding to risky projects and support market entry."[41] To illustrate, Google's chief economist has claimed that Google's acquisitions are primarily "acquihires:" the median number of employees in companies acquired by Google is 6; 25 percent had 3 or fewer employees; 75 percent had 18 or fewer employees.[42] While preventing large companies from hiring additional employees may benefit competitors to those large companies, it is unlikely to help competition or consumers.

The objection suggested by the current paper is that any approach that is conditioned on "bigness" is not tailored to the issues inherent in assessing effects in winner-take-all markets described above. Specifically, "bigness" is clearly not a sufficient indicator of transactions or conduct that may implicate very large effects. Simply put, there exist many "big" companies that do not operate in winner-take-all markets.

Instead, a bifurcated approach could better address the issues described above by focusing the first step on identifying instances where a winner-take-all dynamic is present. If a winner-take-all dynamic is indeed present, then the second step would entail an examination of the evidence in relation to a weaker legal test. For example, the burden could shift to the defendant to demonstrate that the transaction or conduct would not harm competition. Alternatively, the fact finder could judge whether it is *possible* that a transaction or conduct would cause substantial effects.

38 Sandeep Vaheesan, "Unleash the Existing Anti-Monopoly Arsenal: Corporate power can be neutralized if federal agencies simply used the prodigious authority they've been granted," *The American Prospect* (September 24, 2019), available at https://prospect.org/day-one-agenda/unleash-anti-monopoly-arsenal/.

39 University of Chicago, Stigler Center, *Stigler Committee on Digital Platforms,* Final Report (September 2019), available at https://research.chicagobooth.edu/stigler/events/single-events/antitrust-competition-conference/digital-platforms-committee, at pgs. 98-99.

40 University of Chicago, Stigler Center, *Stigler Committee on Digital Platforms,* Final Report (September 2019), available at https://research.chicagobooth.edu/stigler/events/single-events/antitrust-competition-conference/digital-platforms-committee at pg. 99.

41 United Kingdom, *Unlocking Digital Competition: Report of the Digital Competition Expert Panel,* also known as the 'Furman Report' (March 2019) para. 3.102, available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf.

42 Hal Varian, "The Seven Deadly Sins of Tech?" Presentation at the Global Antitrust Economics Conference, New York University, Stern School of Business (May 2019), available at https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=14&ved=2ahUKEwjDweuA4b_lAhVog-AKHakgCQY4ChAWMAN6BAgDEAI&url=https%3A%2F%2Fwww.concurrences.com%2FIMG%2Fpdf%2Fconcurrences_190531_gae_presentations.pdf%3F50755%2F072841c64e7d6daa8f5d7fc2f9d660ef4aa302be&usg=AOvVaw2oh_6nE0vJuP6oNkShBMFR.

VI. CONCLUSION

In prospective cases – that is cases where a new product or service has not yet been proven to be significant – substantial uncertainty is inherent and unavoidable. Even in retrospective cases – that is cases when the product or service has matured – it is frequently very difficult to say whether an acquisition or conduct should have been enjoined. That is, whether outcomes would have been "better" absent the conduct or acquisition. Consider Google's acquisition of Android. It is easy to identify facts about the world in which Google did acquire Android. For example, it is a fact that a large majority of portable devices currently run on Android. But opining whether consumers and competition would be in a better place in a world where Google was stopped from acquiring Android requires knowing how the mobile device marketplace would have evolved absent the acquisition. Such an endeavor, to put it mildly, is not easy.

In this sense, there is likely no silver bullet that will ensure that enforcement decisions are always correct. However, recognizing existing biases in competition law enforcement and trying to minimize them can go some way toward improving outcomes.

Moreover, improving current competition law enforcement is valuable because it provides a tangible response to widespread calls for potentially dramatic change. Calls to regulate big tech companies like we do a municipal utility are indeed dramatic and, we believe, are likely to be harmful on balance. Instead, we believe that a better path forward is to give competition — and antitrust enforcement — a shot.

BIG DATA, BIG TARGET FOR EU ANTITRUST ENFORCEMENT?

BY JAY MODRALL[1]

1 Norton Rose Fulbright LLP, Brussels.

I. INTRODUCTION

The antitrust implications of big data have been a key concern for international antitrust authorities since at least 2016, when the French and German authorities released a study on competition and big data[2] and the OECD organized a roundtable.[3] Since then, antitrust authorities' focus on big data has only increased, with major studies on big data – as well as other digital economy issues – around the globe, from Australia[4] to the United States[5] to the United Kingdom,[6] among many others.

At the European Union ("EU") level, Commission Executive Vice-President Vestager, who recently began a new five-year mandate as Competition Commissioner, has spoken extensively on big data since a major 2016 speech[7] on the topic. Vestager has, however, taken a more nuanced approach than many commentators. In September 2016, she noted that "companies need to make sure they don't use data in a way that stops others competing. But that doesn't mean there's a problem, just because you hold a large amount of data. . . . We don't just assume that holding a large amount of data lets you stop others competing."

More recently, Vestager has sometimes sounded more concerned, commenting that companies "sometimes depend on data, to be able to compete"[8] and that "to keep markets open and competitive, we may need to require companies that control [big] data – such as platforms – to share the data they hold."[9] On the other hand, both in 2016 and her most recent speeches, Vestager has noted that the European Commission ("EC") has examined mergers of companies with important data assets without finding problems due to the combination of the parties' data. "Either the data involved in the merger wasn't that important," she noted, "or similar data was still available to rivals, despite the merger."

So how will the EC enforce EU competition law in relation to big data, now that its new mandate has begun? On the one hand, the EC is rightly afraid of leaping to conclusions, and has examined numerous mergers without identifying specifically big-data-related concerns. On the other hand, like many other authorities, the EC is concerned about the competitive impact of big data and may well feel political pressure to turn up the heat on companies that accumulate and use big data.

This article discusses some specific changes that we may expect to see in the EC's approach to big data issues, drawing in particular from the May 2019 report delivered by three high-level advisors (the "Digital Era Report," or the "Report"),[10] which recommend significant changes to EU antitrust policy, including notably in relation to big data.

Perhaps following Vestager's lead, the Digital Era Report takes a rather nuanced approach to big data. The Report notes that "[t]he significance of data and data access for competition will . . . always depend on an analysis of the specificities of a given market, the type of data, and data usage in a given case." In relation to proposals to require dominant companies to grant competitors access to data under Article 102 of the Treaty on the Functioning of the European Union ("TFEU"), the Report stressed the need for care, noting that "it is necessary to distinguish between different forms of data, levels of data access, and data uses."

---

2 Autorité de la Concurrence and Bundeskartellamt, Competition Law and Data, May 10, 2016, available at https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Big%20Data%20Papier.pdf?__blob=publicationFile&v=2 (the Franco-German Report).

3 See https://www.oecd.org/competition/big-data-bringing-competition-policy-to-the-digital-era.htm.

4 Australian Competition & Consumer Commission, Digital Platforms Inquiry Final Report, June 2019, available at https://www.accc.gov.au/system/files/Digital%20platforms%20inquiry%20-%20final%20report.pdf.

5 Federal Trade Commission, Hearings on Competition and Consumer Protection in the 21st Century, available at https://www.ftc.gov/policy/hearings-competition-consumer-protection.

6 Digital Competition Expert Panel, Unlocking Digital Competition, March 2019, available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf.

7 Margrethe Vestager, Competition and Big Data, September 29, 2016, available at https://wayback.archive-it.org/12090/20191129222113/https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/big-data-and-competition_en.

8 Margrethe Vestager, Defining markets in a new age, December 9, 2019, available at https://ec.europa.eu/commission/commissioners/2019-2024/vestager/announcements/defining-markets-new-age_en.

9 Margrethe Vestager, Digital power at the service of humanity, November 29, 2019, available at https://wayback.archive-it.org/12090/20191130155750/https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/digital-power-service-humanity_en.

10 Jacques Crémer, Yves-Alexandre de Montjoye and Heike Schweitzer, Competition Policy for the Digital Era, available at https://ec.europa.eu/competition/publications/reports/kd0419345enn.pdf.

II. MARKET DEFINITION AND MARKET POWER

Although not specifically in connection with big data, the Digital Era Report argues that the digital economy requires a fundamental re-thinking of the concepts of market definition and market power. "In the digital world," it argues, "it is less clear that we can identify well-defined markets. Furthermore, in the case of platforms, [market] interdependence . . . becomes a crucial part of the analysis whereas the role of market definition traditionally has been to isolate problems. Therefore, in digital markets, less emphasis should be put on the market definition part of the analysis, and more importance attributed to the theories of harm and identification of anti-competitive strategies." The Digital Era Report proposes introducing the concept of a digital "ecosystem" as a framework for analysing anti-competitive effects, particularly in relation to after-markets.

The Digital Era Report also argues for a broadening of the concept of market power, noting that there can be market power even in an apparently fragmented marketplace. The Report notes that market shares are often not a useful concept, particularly in the context of online platforms. In addition to the traditional statistics-based approach to analysing market power, the Report advocates using the concept of intermediation power in relation to online platforms and the power conferred by the collection of data not available to competitors.

III. MANDATING DATA ACCESS AND ARTICLE 102 TFEU

Much of the Report's big-data discussion focuses on conditions in which antitrust law can be used as a tool to require dominant companies to give competitors access to their data. The Report examines arguments for and against mandating data access in different contexts, depending on how the data are obtained (whether they are volunteered, observed, or inferred) and how data are used (use of non-anonymous individual-level data, anonymous individual level data, aggregated data, or contextual data).

The Report notes that the "debate is mostly framed as a debate on whether the criteria of the so-called "essential facilities" doctrine ("EFD") are met." In the data access context, however, the Report argues against application of the EFD and for the application of a broader interest-balancing test. In particular, the Report argues against application of the requirement that a claimant under the EFD show that it requires access to the essential facility to develop a "new product." The goal of the "new product rule" is to preserve a dominant firm's investment incentives by ensuring a sufficient degree of appropriability of profits for the firm. But the Report argues that in some "access to data" cases, the danger that the firm will lose the incentive to invest in data collection and innovation is lower. Thus, rather than a formal "new product rule," authorities applying Article 102 TFEU to "access to data" cases should verify that an access mandate will not eliminate the appropriability of benefits resulting from successful investments. According to the Report, this consideration may be especially important outside the platform context.

As under the "standard" EFD, the Report argues that a precondition for data access remedies under Article 102 TFEU should be indispensability. To determine whether data access is indispensable for Article 102 TFEU purposes, it is important to distinguish access to a data set from access to the data in close to real-time where no substitutes exist. The non-substitutability of data may result from the richness ("number of columns") and size ("number of rows") of the dataset, in particular where machine-learning algorithms play a role. The frequency of data generation also needs to be taken into account when discussing data access, both for providing a service to the person who generated the data and for aggregated applications.

The balancing of interests needs to take account both of the need to protect the dominant firm's investment incentives and the need to ensure that strongly entrenched positions of market power, protected by high barriers of entry, remain contestable. In some settings, such interest-balancing can result in a duty to grant access to data in a form that allows competitors to compete effectively in neighbouring markets, which may include a duty to ensure data interoperability.

The Report outlines several data access "scenarios" to identify situations in which data access mandates would be more or less likely to be appropriate. Access mandates would more likely be appropriate where the data controller holds a gatekeeper position (i.e. access to its data is essential for competing on one or more neighbouring markets) or data access requests are somewhat standardised. Access mandates would be less appropriate where a firm requests data for the purpose of training algorithms for uses unrelated to the fields of activity of the data controller, although even in such case an access mandate could be imposed "where the dominant firm is a large player in data markets and has an infrastructure for data access requests in place."

In the balancing of interest context, the Report discusses whether the threshold for granting access to data should be lowered as compared to access to infrastructure or access to IPR cases, at least where data is produced as a by-product of another activity and incentives

to generate such data will persist irrespective of a possible access mandate, for instance in relation to machine-generated sensor data in the Internet of Things context. While the value of data to a dominant data controller may be taken into account in the original price (and therefore need to be protected under competition law), competition law should not protect an after-market monopoly.

Indeed, the Report supports a more interventionist approach to data access – and possibly data interoperability – for use in complementary markets or after-markets, which are part of the broader "ecosystem" served by the data controller. The EC and EU courts have frequently had occasion to address the relevance of secondary markets and after-markets in the Article 102 TFEU context. But the Report argues that "some of the specificities of data could imply that the competition policy treatment of access to data should be different from that of standard aftermarkets. . . but there is very little economic analysis of these issues, so [their] conclusions should be considered as very preliminary."

In the specific case of individual-level data, the Report discusses the implications of the data portability right under the EU's General Data Protection Regulation ("GDPR"), but notes that data portability in the GDPR does not provide a right to continuous data access or interoperability, but simply a right to receive a copy of some accumulated past data. Thus, it may facilitate a data subject's switching between services but not facilitate multi-homing or the offering of complementary services, which frequently requires continuous and potentially real-time data access. Accordingly, "more demanding" data access regimes than the GDPR could be mandated under Article 102 TFEU for dominant firms, or through sector-specific regulation in particular to open up secondary markets. Tension with the GDPR may also arise where an access mandate involves personal data. The Report notes that the UK and French competition authorities have resolved this tension in past cases by ordering data access on an opt-out basis or permitted access to data only for specified purposes and specified acts of processing.

Thus, in relation to big data, the Report does not support a generally applicable right of access to data under EU antitrust law, but proposes a new analytical framework to make it easier for the EC to impose data access mandates under Article 102 TFEU, especially in relation to after-markets. The Report recommends moving away from the established EFD analysis, applying a more general interest-balancing test.

IV. SHARING AND POOLING DATA AND ARTICLE 101 TFEU

Although much of its discussion focuses on the potential imposition of data-sharing requirements, the Report also discusses the antitrust treatment of voluntary data sharing and pooling. Data sharing and data pooling arrangements will frequently be procompetitive, enabling firms to develop new or better products or services or to train algorithms on a broader, more meaningful basis. Such arrangements can be anti-competitive, however, for example where some competitors are denied access or granted access only on less favourable terms; data sharing amounts to an anti-competitive information exchange including competitively sensitive information; data sharing or pooling discourages competitors from differentiating and improving their own data collection and analytics pipelines; or where the granting of access to data on non- fair, reasonable and non-discriminatory ("FRAND") terms results in an exploitative abuse. Where a dominant, vertically integrated platform provides privileged data access to its own subsidiaries, this can constitute a form of self-preferencing. Where a data pool has market power and gives its members a significant advantage, the pool may be under a duty to give access to others, perhaps on FRAND or similar terms.

The Report references the treatment of data access issues under Article 102 TFEU and merger control, as well as established rules on R&D agreements and patent pools, but notes that it may be more difficult to establish market power of a data pool based on market shares due to the multi-purpose use of data. Thus, an access regime may need to differ depending on the type of use, and the duty to give access should be proportional to the pool's market power. For example, a group of smaller players pooling their data to gain a competitive advantage should not be forced to give their pooled data to a much larger player. Where there is a FRAND or similar duty to provide data access, and the pool's data format standard is proprietary, the standard owner should not be able to raise its fees over time as the pool becomes more important in the market (by analogy to "patent ambushes"), and such duties should apply to both access to the data pool and the use of the data format standard.

The Report notes that the competition law treatment of data sharing and pooling arrangements is relatively new and under-researched, and recommends that the EC use guidance letters and "no infringement" decisions, as well as the next version of the EC's Guidelines on horizontal cooperation. Indeed, a separate block exemption regulation on data sharing and pooling may be called for.

In an open-ended but interesting aside, the Report notes that legislation is required to give firms guidance on what is allowed and not allowed to encourage fuller development of data markets, data sharing and data pooling arrangements. Such a general legal framework would, for example, define contractual rules for access to non-personal data and empower institutions to facilitate the management of consent into the processing of personal data. Such a legal framework would also reduce the pressure on competition law to revolve concerns about access to data.

V. MERGER CONTROL

Perhaps surprisingly, the Report does not discuss the analysis of big data in the merger review context. Rather, the Report focuses on cases in which a dominant platform and/or ecosystem benefitting from strong positive network effects acquires a target with low turnover but a large and/or fast-growing user base and a high future market potential. The competitive concerns arise in markets characterised by high concentration and barriers to entry, resulting from strong positive network effects and possibly reinforced by data-driven feedback loops. The Report notes that this scenario concerns a relatively small group of cases and rejects the common label, "killer acquisitions," noting that the targets in this scenario are typically integrated into the acquirer's ecosystem, not shut down.

Back in 2016, the Franco-German Report expressed concern that a combination of "data troves could raise competition concerns if the combination of data makes it impossible for competitors to replicate the information possessed by the merged entity." As mentioned, however, the EC has investigated this theory a number of times in recent years but have not yet found a cause for concern. The issue will no doubt continue to arise in future cases, but there does not appear to be strong pressure for a change in the EC's current approach.

VI. *QUO VADIS*, EC?

With the Digital Era Report in the EC's arsenal, a new mandate, and political support for scrutiny of big data, what changes can we expect in EU competition law enforcement in relation to big data? As Yogi Berra supposedly said, it's tough to make predictions, especially about the future. But we can make some educated guesses in each of the four areas discussed above.

First, the Digital Era Report recommends that the EC review its approach to defining antitrust markets and assessing market power. Indeed, Vestager recently announced[11] a review of the EC's 1997 notice on market definition.[12] Vestager noted that "digitisation . . . raises some more challenging questions for the way we define product markets. Very often, we find that big digital businesses . . . [provide] consumers with an ecosystem of services, that are all designed to work together well. And as the special advisers pointed out in their report, it can be difficult for consumers to switch from one ecosystem to another." Though the results of the EC's review won't be known until late 2020 or 2021, we can expect the EC to consider a wider perspective on markets that are connected with one another in "ecosystems" and to focus on the role of data in locking customers into such ecosystems.

Second, in relation to the possibility that Article 102 TFEU may require companies to share their data, the EC can be expected to explore new theories of harm, or revisions to existing theories on a case-by-case basis. Such mandates would only be imposed on companies found to have market power, as is the case today, but changes to the EC's traditional assessment of market definition and market power may make findings of market power more frequent. Moreover, if the EC follows the Digital Era Report's recommendation to relax the traditional EFD criteria, the EC may be more likely to impose data access remedies in the coming years. If so, companies active in complementary or downstream markets are more likely to benefit than head-to-head competitors of the dataholder. On the other hand, any such changes to Article 102 TFEU enforcement in relation to big data will emerge, if at all, only through individual cases, which are few and far between and normally take years to resolve.

Third, in relation to the treatment of data sharing and pooling under Article 101 TFEU, the Digital Era Report called on the EC to provide additional guidance. In fact, in late 2019, the EC launched a review of its horizontal block exemption regulations and guidance on horizontal co-operation, including notably sharing of competitively sensitive information.[13] This review was triggered by the upcoming expiration of the relevant block exemptions, not by the Digital Era Report recommendations, but it seems likely that the EC will take this opportunity to clarify the treatment of data sharing and pooling. A separate block exemption on data sharing and pooling, as proposed in the Report, could be very helpful.

Fourth, merger control seems to be the area in which major changes in relation to big data are least likely. The EC has developed a methodology for analyzing big data combinations in recent years, and there seems to be no strong pressure for change. Each case is different, but the EC's practice so far suggests that the combination of merging parties' "data troves," without more, are unlikely to raise concerns.

---

11 Margrethe Vestager, Defining markets in a new age, December 9, 2019, available at https://ec.europa.eu/commission/commissioners/2019-2024/vestager/announcements/defining-markets-new-age_en.

12 Commission Notice on the definition of relevant market for the purposes of Community competition law, OJ C 372, 9.12.97, pp 5-13.

13 See https://ec.europa.eu/competition/consultations/2019_hbers/index_en.html.

In summary, in the new EC mandate we are likely to see an increased focus on big data in EU competition law enforcement. A number of independent initiatives already under way could facilitate change. However, any such changes are likely to be incremental. A senior EC official recently told me that there are more conferences on big data than cases. That may continue to be true.

On the other hand, competition law enforcement is not the only game in town. Indeed, Executive Vice-President Vestager observed in September 2019 that "data is not just an issue for competition. . . So we may also need broader rules to make sure that the way companies collect and use data doesn't harm the fundamental values of our society."[14] Indeed, Commission President von der Leyen has charged Vestager – in her parallel capacity as Commissioner for making Europe fit for the digital age -- with co-ordinating work on new legislation on artificial intelligence and digital services. The scope of such legislation is not yet clear, but ultimately Vice-President Vestager may have a greater impact on the collection and use of big data in Europe in her regulatory capacity than in her antitrust enforcer capacity.

14 Margrethe Vestager, Security and trust in a digital world, September 13, 2019, available at https://wayback.archive-it.org/12090/20191130063011/https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/security-and-trust-digital-world_en.

CAN DIGITAL DATA BE REPLACED? DATA SUBSTITUTABILITY IS THE KEY

BY MARIATERESA MAGGIOLINO[1] & GIULIA FERRARI[2]

1 Associate Professor, Department of Law, and Director of the Master of Arts in Law, Bocconi University.

2 PhD Candidate in Legal Studies, Bocconi University.

I. INTRODUCTION

"Just the facts ma'am: competition cases are all about facts."[3] That sentence has the merit of highlighting how antitrust law, aiming at ensuring the proper functioning of the market, cannot disregard the knowledge of the many and varied empirical elements that influence economic agents' decisions and their effects on consumer welfare. This is the reason why, for example, antitrust authorities and judges: (i) investigate the way the demand is oriented in order to understand consumers' preferences and spending constraints; (ii) identify the business models used by undertakings to grasp the economic rationality underlying their strategies; and, finally, (iii) describe the structural and institutional characteristics of markets in order to establish the empirical elements that affect businesses and consumers' choices.

Therefore, there cannot be a "one-size-fits-all" answer to the question of whether data constitute an essential resource (i.e. facility) or a barrier to entry. This question cannot be answered without taking into consideration the specific circumstances that characterize the particular case under scrutiny.

What instead does not vary depending on the facts of the case at stake is the analytical framework that those in the antitrust field use to develop their reasoning. Thus, let's try to reconstruct it.

In essence, when we say that a resource used in a given market to produce the goods there sold works as a barrier sheltering that market, we mean that an undertaking that wants to provide those goods must incur the costs necessary to use that resource: the higher the costs, the higher the barrier. With the same degree of approximation, when we say that a resource is essential, we mean both that, without that resource, a company will not be able to create its products and/or services (obviously to sell them in a given market), and that the said resource does not have economically reproducible substitutes. From this point of view, therefore, it could also be argued that an essential resource is an insurmountable barrier to entry. But this is not the point.

The point is that, in both cases, the interpreter's main task will be to understand whether and to what extent the resource in question has viable substitutes, which could possibly be cheaper. In the event that those substitutes are deemed to exist, the resource could certainly not be considered to be essential; it could be considered, at most, as a barrier to entry, depending on the cost of the available substitutes.

By applying this reasoning in relation to data, the crucial issue is thus to understand if and to what extent digital data have substitutes. In other words, what matters the most in the context of this competitive assessment is the existence (or not) of other "elements" (and mainly other data) that can be considered to be adequate substitutes for the resource in question, i.e. digital data.

II. DATA SUBSTITUTABILITY

Reading about the digital economy, it is common to come across the observation that digital data are non-rivalrous resources. The term refers to the idea – which is certainly shared here – that the use by Undertaking A of certain data would not prevent Undertaking B from using the same data at the same time and equally benefitting from them, since, like other intangible assets, digital data do not lose value with use.[4]

In addition, the Fourth Industrial Revolution has chosen binary code as the language for the representation of the world and human behaviors. As a result, many objects have become sources of digital data: not only personal computers and the Internet but also devices – i.e. intelligent objects – in everyday use and of modest cost.[5] Since nothing leads to the belief that this tendency towards the "dataization" of reality is likely to be interrupted, it is easy to understand why the concept of ubiquity is often attributed to digital data.[6] In the context of the digital economy, that expression refers to the fact that potentially all companies are in a position not only to generate, but also to receive and retain huge volumes of

---

3 Ian Forrester, *A Bush in Need of Pruning: the Luxuriant Growth of Light Judicial Review* 410, Claus-Dieter Ehlermann, Mel Marquis, *European Competition Law Annual 2009: Evaluation of Evidence and Its Judicial Review in Competition Cases* (2011).

4 As evidence of this, it is noted that undertakings interested in using digital data may choose to acquire such resources from the market. Datasets made available on the market, either free of charge or for a fee, may simultaneously be the subject of several contracts and thus multiple uses by different undertakings without reducing their value, i.e. the possibility for undertakings "buying" such data to use them profitably. See Anja Lambrecht & Catherine Tucker, *Can Big Data Protect a Firm from Competition?* 4-5 (2015), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2705530, and Alessandro Acquisti & Hal Varian, *Conditioning Prices on Purchase History*, 24 Marketing Sci. 367 (2005).

5 Reference is made to the wide variety of "smart devices" (such as virtual assistants based on AI technology) and to the so-called Internet of Things (IoT).

6 See Catherine Tucker, *The Implications of improved attribution and measurability for antitrust and privacy in online advertising markets*, 20 *Geo. Mason L. Rev.* 1025, 1030 (2013) and Manuel Castells, *La nascita della società in rete* (2014).

data. It is thus easy to understand why many scholars argue that companies that are interested in exploiting these data only must face the fixed costs of acquiring the infrastructure and technical skills needed to generate and collect them.[7]

However, these general statements, while correct and acceptable, do not help competition law scholars and practitioners to understand whether, in a specific case, a certain set of data could be considered substitutable with others.

In competition law, indeed, the substitutability of any good, be it an input or a product (whether intermediate or final) must be assessed in relative – rather than absolute – terms. The assessment exercise must be accomplished by understanding the use that is made of that particular resource, which in turn can be identified based on the needs and desires that it satisfies. Therefore, in order to understand whether data – or, *rectius*, datasets – have substitutes or not, it is essential to understand what those data are intended for.

Suppose, for example, that a businesswoman active in the tourism sector of the Riviera Romagnola in Italy wants to know the eating habits of her potential customers, as this information could enable her to personalize the menus of the chain of hotels and guesthouses she manages. To this end, already equipped with the software and skills necessary to develop inferential analysis, she strives to obtain as many useful data as possible on the eating habits of those visiting Romagna so as to map in the most accurate way the target she wishes to please with her services. To do this she probably intends to merge those data with those that she already has, having presumably collected them during the years of activity of her hotels and/or obtained from her trade association.

In order to obtain such data, she can, for instance, turn to Google Search, since the American company certainly has traces of the food and wine searches made by users who, during the summer, connected to the Internet from Romagna. Similarly, she may seek to obtain the data that platforms such as TripAdvisor and The Fork certainly have collected when their users consulted them between May and September to find out about the restaurant and food services available in the Romagna provinces. Furthermore, the businesswoman can address insurance companies, companies offering navigation systems and/or car rentals in order to find out which restaurants and culinary establishments those who travelled the roads of Romagna in the summer visited.

In short, since there are several sources of data capable of revealing the information required (in our example the eating habits of those who visit the Romagna region), all the datasets above mentioned must be considered as belonging to the same market – i.e. the market of the data from which the eating habits of customers on the Riviera Romagnola are inferred – and, thus, all the firms offering them must consider themselves to be competitors for the supply of such data.

Many other similar examples can be made: let's think for instance of the reading, music and movie preferences of a particular group of users (such as, young people between 20 and 30 years old living in Europe) that some undertakings may want to know in order to develop successful "products" (such as, reading apps, music and streaming apps, but also songs and movies/tv shows specifically targeting the preferences of this segment of the public). This information may be available both to search engine providers by way of analyzing the search queries that these users have entered over a given period of time, while a social network provider may also be able to gain the same knowledge by looking at the profile information that users have shared on its platforms. In addition, the same information may be held, respectively, by platforms offering reading services (such as, to name a few, Blinkist, Instaread, or Amazon Kindle), music streaming platforms (such as Spotify or Apple Music), as well as streaming platforms (such as Netflix, Sky Go, or Amazon Prime Video).

Again by way of example, an undertaking wishing to develop a word recognition program based on artificial intelligence ("AI") technology would need to have digital versions of texts – if necessary translated into several languages – for its machines to read, so that it can design and test the algorithm which, over time, will learn to identify words, assign them meanings and translate them.[8] Even in this scenario, the undertaking may get access to such texts through a contract with libraries, but those texts can be also replaced by other texts held by the same or other libraries.

Therefore, depending on the particular information at issue, all the said companies/libraries can be considered to be competitors as they may all be in the position to "answer" the same needs and uses. In each of the examples above, indeed, the competitive analysis should aim

---

7 Carl Shapiro & Hal R. Varian, *Information rules: A strategic guide to the network economy* 24 (1999).

8 This is for instance the case of the Google Books project: through a set of "digitalization agreements" entered into with a number of libraries starting from 2002 the US company was able to obtain a digitalized copy of the books included in their catalogues. By doing so Google was also able to achieve at least two different objectives: (i) refine its search and data mining activity within the digitized texts and (ii) develop further products and services based on that data and the information extracted from it, such as the translation service Google Translate.

at understanding if and how much the resources in question (data about the eating habits, reading/music/movie preferences or texts) can be replaced by other data held by different subjects.

Out of these examples, it must be noted that typically any information can be inferred by querying different and distinct datasets, not necessarily controlled by the same subject. Since digital data are nothing more than the representation in sequences of ones and zeroes of real-world facts and related human behaviors, and since, due to the Fourth Industrial Revolution, digital data can be obtained either by transforming analog data into digital data, or gathering the inputs given by users not only on the Internet or the IT structures of companies, but also offline through intelligent objects, information such as the eating habits of those who visit the shores of Romagna (or the reading/music/movie preferences of young European or texts) can be inferred from a number of various sources. Moreover, since under antitrust law, all goods responding to the same needs and satisfying the same desires are deemed to belong to the same market, then all these datasets must be deemed to belong to the same market.

Of course, going back to our examples, it could be argued that not all those datasets are perfectly equivalent to each other in terms of quantity, quality and accuracy. Nevertheless, in the everyday life of market analysts, it rarely happens that products and services belonging to the same market are *perfectly* interchangeable. A certain degree of heterogeneity is indeed inherent in the contemporary economy, especially in all the markets that are not oligopolistic.

Moreover, it is doubtful that the hotelier from Romagna or the company wishing to develop or ameliorate new products/services targeting young European or AI software would be refused almost all the requests made to the subjects owning the desired data. This possibility (i.e. the refusal to provide the data) – which often attracts the attention of commentators whose intent is to (uncritically) discuss the indispensable nature of "big data" – sets the way, though, to discuss another important question, i.e. the access to data. Discussing of data access, however, reinforces the importance - as discussed so far - of analyzing the factual elements that serve to identify the relevant markets.

In other words, stating that several datasets similarly capable of generating information belong to the same relevant market is a separate (and logically prior) matter to the question of whether access to some of those datasets is prevented by the companies that control them. Indeed, while the first consideration concerns the description of the markets, the second concerns the behavior of the undertakings. Specifically, the latter concerns a possible refusal to deal which the antitrust authorities would be able to pursue – whether successfully or not is still a different issue – if it resulted from a concerted action or, more likely, from a unilateral act by an undertaking holding a dominant position.

Finally, it could be argued that the decision to identify markets for data on the basis of the use that is made of them does not allow *per se* to achieve a result that some commentators hope to reach: namely having a perfect coincidence between the resource named "dataset α" and the relevant market. This coincidence is what in fact would allow antitrust law enforcement agencies to nimbly show the dominance of a given company on the market for the sale of data α.

In this regard, it is necessary to focus on a cornerstone of antitrust analysis which can be recalled as follows: it may happen that a resource constitutes its own market, i.e. that a resource does not have substitutes. Nevertheless, this possibility – which may arise more or less frequently – must be ascertained on a case-by-case basis, i.e. after having assessed and ruled out the existence of possible substitutes based on factual elements. In more sophisticated terms, the possibility that a resource constitutes a market in and of itself is an issue that may prove empirically grounded in an individual case, but conceptually is not always true.

It is not by chance that opposite scenarios are found in the decision-making practice.

For example, the European Commission chose to clear the *Facebook-WhatsApp* merger as the empirical analysis allowed the authority to verify that despite the merger between the datasets owned by the two companies, the resulting datasets continued to have substitutes on the market; in other words, the Commission ruled out that Facebook would have become the sole gatekeeper of users' digital data as a result of the acquisition of WhatsApp.

Indeed, the Commission clearly stated that, even after the transaction, "there will continue to be a large amount of Internet user data that are valuable for advertising purposes and that are not within Facebook's exclusive control"[9]. This supported the argument that, post-merger, the

---

9 European Commission, 3 October 2014, case COMP/M.7217, §§ 188-189. See also the case *TomTom/Tele Atlas* (European Commission, 14 May 2008, case COMP/M.4854) and the case *Google/DoubleClick* (European Commission, 11 March 2008, case COMP/M.4731) for other hypothesis in which the Commission denied the occurrence of an anticompetitive foreclosure.

parties' competitors, including telephone companies and other digital platforms, were still able to access alternative sources of commercially useful Internet users data.

On the contrary, in the recent Italian *Enel* and *Acea* cases,[10] the Italian Antitrust Authority (Autorità Garante della Concorrenza e del Mercato) ascertained the non-substitutable nature of the lists of customers to which the two companies had exclusive access. The lists contained the name of the clients served by the two energy companies on the regulated market, who had given their (privacy) consent to be contacted for commercial purposes for the supply of liberalized energy services.[11]

Specifically, the Authority maintained that such lists of customers could not be replicated by other companies that, unlike Enel and Acea, were not vertically integrated and thus could not serve customers on the regulated market.[12] It consequently concluded that the exclusive control of these customers' data by Enel and Acea could produce an abusive form of anti-competitive foreclosure.[13]

III. CONCLUSIONS

While this article does not set out to *a priori* exclude that the datasets held by certain online platforms (such as the so-called GAFA[14] companies) could give rise to a competitive advantage for incumbents and an entry barrier for potential competitors due to the quantity, quality and accuracy of the data concerned, we want to highlight that an empirical analysis of the factual circumstances on a case-by-case basis should always be carried out to assess whether the same data (i.e. the data responding to the same needs) could be obtained elsewhere on the market.

This empirical analysis, indeed, is the first and essential step to establish – on the basis of solid grounds – whether such datasets in fact constitute an entry barrier and, if so, the "height" of any such barrier. At the same time, the analysis of the same circumstances will be crucial to identify whether the datasets can be deemed to amount to an essential resource. This is the case, of course, where it is ascertained that the data in question do not have any plausible substitutes. In this scenario, then, the task of the antitrust interpreters would be that of assessing the existence of the behavioral elements of the violation, i.e. a refusal to access the essential resource, should the other building blocks of the antitrust breach be deemed to arise – whether in the form of a concerted practice or an abuse of dominance.

To this end, data substitutability should always be assessed in relation to the kind of information that the subject(s) who seek(s) it want(s) to infer from those data. Lacking a "one-size-fits-all" answer to data substitutability, a correct analysis of whether or not certain data have substitutes can only take place after having thoughtfully established the very needs and uses that the "information seeker" wants to accomplish through such data. Assessing these preferences is what helps the antitrust interpreter to identify the kind of information an undertaking wants to infer from the data it seeks, and thus their potential substitutes.

However, there is a "but."

While data substitutability is key to establish whether data constitute a barrier to entry or an essential resource, another different data-related issue arises as a specific aspect of the Fourth Industrial Revolution. Big data companies do not always know in advance the kind of information they want to seek by analyzing data, but they will nevertheless find it very rewarding to analyze large volumes of different datasets to

---

10 Autorità Garante della Concorrenza e del Mercato (AGCM), Decision no. 27494 of 20 December 2018, case A511 – *ENEL - condotte anticoncorrenziali nel mercato della vendita di energia elettrica*, and Decision no. 27496 of 20 December 2018 - *A513 – ACEA - condotte anticoncorrenziali nel mercato della vendita di energia elettrica*. It should be noted that the paper focused on the *Enel* case as it is considered to be more explanatory of the critical competitive issues related to the collection and use of data, and because this decision has passed the Lazio Regional Administrative Court's scrutiny, which instead annulled the AGCM's Decision of the *Acea* case. On the other hand, the Regional Administrative Tribunal ("TAR") has imposed a recalculation of the fine inflicted on Enel on the basis of the parameters identified by the Administrative Judge in the ruling. See Tar Lazio, Judgment no. 11976/2019 of 17 October 2019 (*Acea* case) and Tar Lazio, Judgment No. 11958/2019 of 17 October 2019 (*Enel* case).

11 Pursuant to Law 124/2017 the Italian electricity market should have been fully liberalized as of July 1, 2020 when the regulated regime of "greater protection" (according to which domestic and small users can be provided with services at a regulated price) would have been totally replaced by the free market. However, in December 2019 the Italian Parliament passed an amendment according to which the full liberation will take place starting from January 2022.

12 Indeed, only the companies acting as distributors of electricity can provide the regulated energy services. See §§ 87 and 226 of the Decision in the *Enel* case.

13 More in detail, Enel has managed to acquire such data by offering their customer base the possibility to provide their consent to the processing of their personal data for commercial and marketing purposes separately to the companies of the Enel group and to third parties, although there was no obligation in this regard under the legislation on the protection of personal data. As a result of such setting, it was found that on average 70 percent of customers gave their consent to be contacted exclusively by companies of the Enel group, while only the remaining 30 percent also gave their consent to the processing of their data by third parties.

14 The acronym refers to the tech companies Google, Amazon, Facebook, and Apple.

hunt for useful information. This process may indeed grant the undertakings with access to such massive amount of data (such as GAFA) an incomparable competitive advantage in that they may be able to find new business opportunities before and better than their competitors. As such, they may discern the opportunity to profitably expand to a collateral industry, and even be able to shorten learning times and thus understand not only in which markets it is worth investing, but also what to do to quickly develop a new business activity.[15]

In such circumstances, however, there is no need to assess whether these data constitute a barrier to entry or an essential resource and, as a consequence, assessing data substitutability cannot add much to the antitrust analysis, mainly because in such a case all data are equally important. In other words, one of the competitive dangers inherent in the Fourth Industrial Revolution relates to the possibility that many undertakings may not be able to form and use – for purposes *ex ante* unknown, but that only the analysis of such data will reveal – the knowledge of the world and of human behaviors that is already available to a few companies, and which grants them a precious competitive advantage.

That said, contemporary antitrust law can do little to face this competitive risk because it assumes by its very nature a position of partial economic equilibrium, as it proceeds "market by market," ascertaining the power that a company enjoys in a given market segment and verifying that the conduct of that company does not worsen, to the detriment of consumers, the balance achieved between supply and demand in that same market.[16] Indeed, because of the flexibility of digital firms, contemporary competitive analysis should instead focus on the economic power of firms and the effects of their behavior on several markets at the same time. At the moment, however, antitrust law interpreters have not yet succeeded in conceiving new categories of analysis that would be useful for this purpose. In order to overcome the problem of equal opportunities with regard to the knowledge of the world and human behaviors, a regulatory solution is seen so far as the most suitable, that is, the opening up of all data by whoever controls them, so that all companies can start from this common basis to develop products and services.[17]

---

15 This is the case of undertakings that have from the outset an ongoing, vast and varied set of updated data.

16 And not for other reasons, such as the fact that antitrust law does not protect the proper functioning of the market, instead of a fair distribution of wealth or other values of a more political nature, nor because it does not establish automatisms, such as that according to which an undertaking in a dominant position which infringes data or consumer protection rules should, for that very reason, be considered as abusing its dominant position within the meaning of 102 TFEU.

17 On this point, see Michal Gal & Daniel Rubinfeld, *Data Standardization*, *NYU Law and Economics Research Paper*, No. 19-17 (2019), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3326377, as well as the contribution published on the website of the European Data Portal (portal established under Directive 2003/98/EC on the re-use of public sector information), *AI and Open Data: a crucial combination*, 4 July 2018, available at the following link: https://www.europeandataportal.eu/en/highlights/ai-and-open-data-crucial-combination.

CPI Subscriptions

CPI reaches more than 35,000 readers in over 150 countries every day. Our online library houses over 23,000 papers, articles and interviews.

Visit competitionpolicyinternational.com today to see our available plans and join CPI's global community of antitrust experts.

CPI | COMPETITION POLICY INTERNATIONAL